16(S4)

HONG KONG MEDICAL JOURNALS

香港醫學雜誌

Volume 16 ▪ Number 5 ▪ October 2010

Supplement 4

HONG KONG SUPPLEMENT 4
VOLUME 16 ▪ NUMBER 5 ▪ OCTOBER 2010

HONG KONG MEDICAL JOURNAL
香港醫學雜誌

The official publication of the
Hong Kong Academy of Medicine
and the Hong Kong Medical Association

Research Fund for the Control of Infectious Diseases

Research Dissemination Reports

控制傳染病研究基金

研究成果報告

Gastrointestinal Diseases
腸胃病

Respiratory Infectious Diseases
呼吸道感染疾病

香港醫學專科學院出版社
HONG KONG ACADEMY OF MEDICINE PRESS

HONG KONG MEDICAL JOURNAL
香港醫學雜誌

Vol 16 No 5 October 2010
Supplement 4

Research Fund for the Control of Infectious Diseases

Research Dissemination Reports

Dissemination reports are concise informative reports of health-related research supported by funds administered by the Food and Health Bureau, namely the *Research Fund for the Control of Infectious Diseases*, the *Health and Health Services Research Fund*, the *Health Care and Promotion Fund* and the *Health Services Research Fund*. In this edition, 12 dissemination reports of funded projects related to gastrointestinal diseases and respiratory infectious diseases are presented. Three of the reports are highlighted, owing to their potentially significant findings, impact on health care delivery and practice, and/or contribution to health policy formulation in Hong Kong.

Human noroviruses (hNoVs) are the leading cause of viral gastroenteritis around the world. Knowledge of the prevalent viral genotypes responsible for sporadic cases and larger outbreaks would be an important epidemiological tool in helping to control its spread. Three major hNoV databases currently exist. However, two are not publicly accessible and the third mainly reports cases occurring in the United Kingdom. Leung et al[1] set out to address this deficiency by creating a publicly accessible norovirus database containing the complete genomes of 31 norovirus isolates obtained from cases in Hong Kong. NoroBase is hosted by the Chinese University of Hong Kong. Users can submit query nucleotide sequences and conduct sequence similarity searches. Currently 80 complete and partial genome sequences have been uploaded. It is hoped that virus genotyping via homology search against NoroBase may benefit local epidemiologists and public health professionals in their studies of the epidemiologic pattern of hNoV infections in Hong Kong and in the identification of aetiologic agents in acute gastroenteritis outbreaks.

During the SARS outbreak in Hong Kong in 2003, many individuals were exposed to the SARS-coronavirus (SARS-CoV) but not all of them became infected. Susceptibility and resistance to infection depends in part on host genetic factors. Two studies in this supplement revealed more about the immunogenetics of SARS infection. Ng et al[2] looked at the differential activation of T cells following exposure to SARS-CoV. A significantly higher frequency of HLA allele DRB4*01010101 was found in the SARS-susceptible than in the SARS-resistant group. In contrast, significantly higher frequencies of HLA-B*1502 and HLADRB3*030101 were found in the SARS-resistant than in the SARS-susceptible group. However, none of these associations was significant after statistical correction. These findings do not suggest a strong involvement of HLA with genetic susceptibility to SARS based on the HLA genes studied.

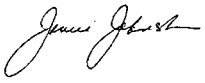
Chan et al[3] examined the potential involvement of the cell adhesion molecule and pathogen recognition receptor, DC-SIGN, as a facilitator of SARS-CoV infection. The investigators hypothesised that particular single nucleotide polymorphisms in the DC-SIGN promoter region may be associated with the severity of SARS-CoV infection. They found that the SNP-336A/G is associated with the levels of serum lactate dehydrogenase on admission, which is an independent prognostic indicator for the severity of SARS. This functional SNP affects the promoter activity of DC-SIGN and may alter gene expression and hence host immune response. In both studies, further evaluation is needed to determine the validity and clinical relevance of their findings.

We hope you will enjoy this selection of research dissemination reports. Electronic copies can be downloaded from the Research Fund Secretariat website (**http://www.fhb.gov.hk/grants**). Researchers interested in the funds administered by the Food and Health Bureau may also visit the website for detailed information about application procedures.

Supplement co-editors

Dr Stewart Chan
Consultant (Research Office)
Food and Health Bureau

Dr Richard A Collins
Scientific Review Director
(Research Office)
Food and Health Bureau



Dr Janice M Johnston
Consultant (Research Office)
Food and Health Bureau

References

1. Leung WK, Chan HL, Lai RW, Sung JJ. Establishment of public norovirus genome database in Hong Kong. Hong Kong Med J 2010;16(Suppl 4):4-7.
2. Ng MH, Cheng SH, Lau KM, et al. Immunogenetics in SARS: a case-control study. Hong Kong Med J 2010;16(Suppl 4):29-33.
3. Chan KY, Xu MS, Ching JC, et al. Association of a single nucleotide polymorphism in the CD209 (DC-SIGN) promoter with SARS severity. Hong Kong Med J 2010;16(Suppl 4):37-42.

WK Leung 梁偉強
HLY Chan 陳力元
RWM Lai 賴偉文
JJY Sung 沈祖堯

Establishment of public orovirus genome database in Hong Kong

Key Messages

1. Phylogenetic analysis demonstrated that a panel of genetically diverse human norovirus (hNoV) strains are circulating in Hong Kong; those belonging to the GII.4 cluster appear to be most prevalent.
2. NoroBase is a publicly accessible web-based hNoV database and provides complete viral genomic and partial gene sequences to facilitate local studies.
3. With the annotated sequences in NoroBase, virus genotyping via homology search against our database provides an efficient way for epidemiologists and public health professionals to study epidemiologic patterns of local hNoV infections and identify and characterise aetiologic agents during gastroenteritis outbreaks.

Introduction

Human norovirus (hNoV) is a leading cause of viral gastroenteritis worldwide, affecting all age-groups in both developing and developed countries.[1,2] Most cases of viral gastroenteritis, both sporadic and outbreak settings, are caused by hNoV.[3-5] The virus was first identified in 1972 and belongs to the genus Norovirus in the family Caliciviridae (previously known as Norwalk-like virus or small round structured virus). It is listed by the National Institute of Allergy and Infectious Diseases of the United States as a Category B priority pathogen. The genome of hNoV consists of a single positive-sense, single-stranded 7.5kb long linear RNA molecule. Currently, hNoV is classified into three genogroups (GI, GII, and GIV), which are sub-divided into more than 20 genotypes such as genogroup II genotype 4 (GII.4).[6] It is highly contagious and known to transmit primarily from person-to-person through the faecal-oral route. An inoculum of no more than 100 viral particles may be sufficient to infect a susceptible individual. Currently no antivirals or vaccines against hNoV are available. Studies on the pathogenesis and vaccine development of hNoV disease are severely hampered by the absence of a robust in vitro culture system or a suitable small animal model. Murine norovirus can replicate in dendritic and macrophage cell lines.[7] There is also evidence suggesting that naturally occurring virulent recombinant norovirus is emerging and circulating in the wild.[8] The high prevalence and mutation rates, the low infectious dose, and the lack of a vaccine highlight the clinical and public health importance of hNoV.

There are three major hNoV databases: CaliciNet (Centers for Disease Control and Prevention in the United States), Norovirus Database (Foodborne Viruses in Europe Network), and Norovirus Molecular Epidemiology Database (Health Protection Agency in the United Kingdom). However, the former two are not open to the public and the latter mainly details strains collected only in the United Kingdom. A database that better represents local circulating strains is therefore desirable. We describe the establishment of a local hNoV nucleotide sequence database that may facilitate epidemiologic investigations of hNoV infections in Hong Kong.

Methods

From December 2004 to November 2006, 687 faecal specimens were collected from patients presenting with acute gastroenteritis symptoms at the Prince of Wales Hospital, Hong Kong. Clinical symptoms included diarrhoea, vomiting, and abdominal pain. Diarrhoea was defined as having more than three loose stools per day. The specimens were stored at -80ºC immediately upon collection. In addition, 30 hNoV-containing faecal specimens collected locally from 2002 to 2004 were provided by Dr WL Lim of Virology Division, Public Health Laboratory Centre, Department of Health, Hong Kong.

Viral RNA was extracted from faecal specimens using QIAamp Viral RNA Mini Kit (Qiagen, US) as per manufacturer's instructions. Extracted RNA was then reverse transcribed in cDNA. Genogroup-specific real-time quantitative polymerase chain reaction (RT-qPCR) assay based on TaqMan technology was used to detect hNoV. Phylogenetic analysis of local hNoV strains was performed based on the N-terminal/Shell domain region of noroviral protein 1 gene. The complete noroviral genome was sequenced by primer walking approach.

Hong Kong Med J 2010;16(Suppl 4):S4-7

The Chinese University of Hong Kong:
Department of Medicine and Therapeutics
WK Leung, HLY Chan, JJY Sung
Department of Microbiology
RWM Lai

RFCID project number: 03040482

Principal applicant and corresponding author:
Dr Wai Keung Leung
Department of Medicine and Therapeutics,
The Chinese University of Hong Kong,
Shatin, NT, Hong Kong SAR, China
Tel: (852) 2632 3173
Fax: (852) 2637 3852
Email: dr_wkleung@alumni.cuhk.net

Results

Norovirus collection and phylogenetic analysis

Using genogroup-specific RT-qPCR, hNoVs were detected in 90 (13.1%) of the 687 faecal specimens. Among these 90 specimens, eight (9%) were infected with GI, 73 (81%) with GII, and nine (10%) with both strains. Of the 90 hNoV, 79 (7 GI and 72 GII) were successfully sequenced for phylogenetic analysis based on the N/S domain region of the VP1 gene. Regarding hNoV GI, it covered five genotypes (GI.3, GI.5, GI.6, GI.7, and GI.8), but no circulating strain predominated. For hNoV GII, it covered eight known genotypes (GII.2, GII.3, GII.4, GII.6, GII.13, GII.14, and GII.16) and one previously undescribed genotype. GII.4 was the most frequently encountered genotype, accounting for 79% of GII strains analysed, followed by GII.3 and GII.6 strains (Fig 1).

Of the 30 archived hNoV-containing faecal specimens, 22 were successfully sequenced for phylogenetic analysis based on the N/S domain region of the VP1 gene. All hNoVs belonged to GII and covered five genotypes (GII.1, GII.2, GII.3, GII.4, and GII.14). GII.4 was the most common genotype (n=13, 59%), followed by GII.1 strain.

Complete genome sequencing

Of the 101 (79+22) hNoV genotyped above, 31 (31%) were successfully sequenced for the complete genome using the primer walking approach. All complete genomes were derived from hNoV GII.4 strains collected at the Prince of Wales Hospital from 2004 through 2006.

Database description and functionality

Complete viral genomic and partial gene sequences can be publicly accessed at the NoroBase (http://norovirus.mect.cuhk.edu.hk or http://www.norovirus.hk). NoroBase is hosted at The Chinese University of Hong Kong. No subscription fee or access authentication is required. Users can submit their query nucleotide sequences in FASTA format. Homology search for closely related local hNoVs is accomplished by NCBI's BLAST program installed on the local server. In the main BLAST result panel of NoroBase, returned local hNoVs are displayed in descending order of hit score. In addition to the E value, and percentages of queries covered and maximum identity users are familiar with on NCBI's online BLAST, NoroBase makes available further information necessary and helpful for epidemiologic investigation, including the year and month of specimen collection, whether the strain was collected from a sporadic case or outbreak settings, and genogroup and genotype clustering (Fig 2).

Discussion

In our study, the proportion of patients with acute gastroenteritis attributed to hNoV infection was 13%. This finding concurs with a recent local epidemiologic study reporting a similar detection rate of about 16% in 2005.[3]

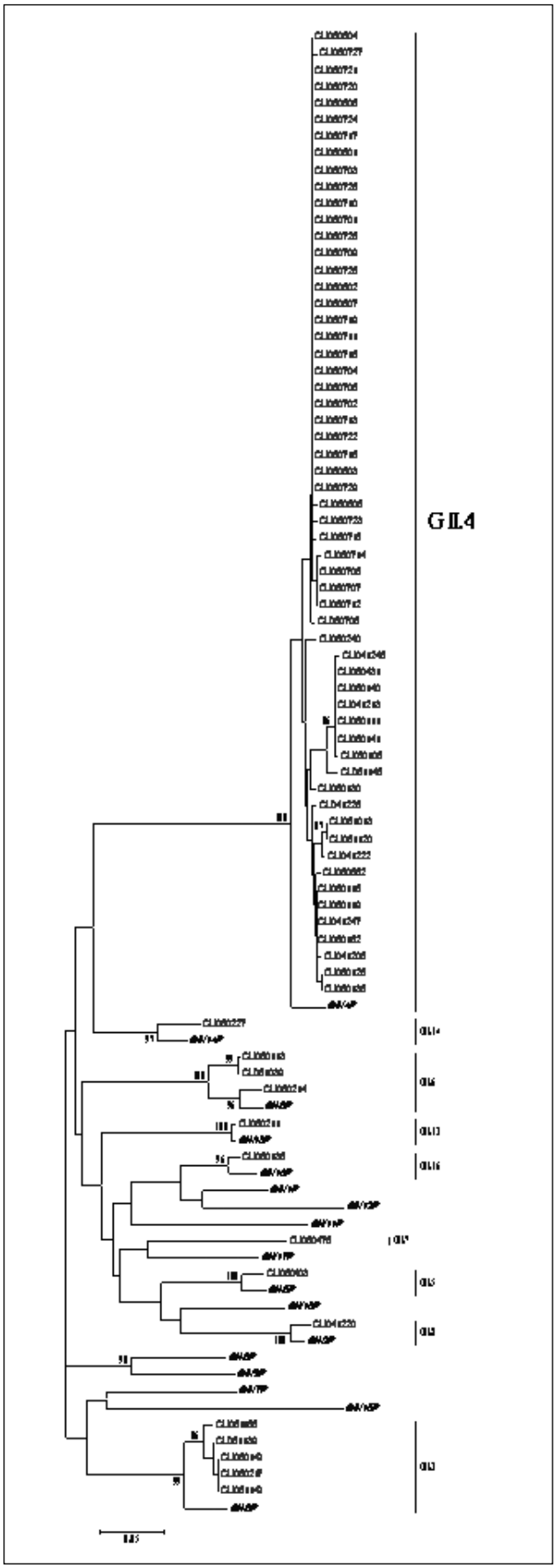


Fig 1. Phylogenetic analysis of human norovirus genogroup II (hNoV GII) strains collected from December 2004 to November 2006 at the Prince of Wales Hospital

Fig 2. Main BLAST result panel of NoroBase

Phylogenetic analysis demonstrated that a panel of genetically diverse hNoV strains that covered two genogroups and 13 genotypes were circulating, with those belonging to GII.4 cluster being the most frequent. This concurs with other epidemiologic studies worldwide. Epidemiologic studies elsewhere have suggested that non-GII.4 strains may more often contribute to institutional outbreaks, foodborne and waterborne transmissions. For instance, in an outbreak of noroviral gastroenteritis affecting more than 1000 evacuees from Hurricane Katrina in the US, multiple strains (GII.2, GII.6, and GII.17) were co-circulating.[9] Coexistence of multiple hNoV strains other than GII.4 was also frequently reported in patients with shellfish consumption as the presumed source of infection.[10] In a waterborne outbreak of viral gastroenteritis, two hNoV stains (GI.3 and GII.6) were detected.[11] Furthermore, strains other than GII.4 have been shown to be endemic in certain countries. In Brazil, a surveillance study demonstrated that GI strains were linked to nearly half of all hNoV infections throughout the year, without there being any marked seasonality.[12] As information on the source of infections was not available in our study, whether co-circulation of multiple strains in Hong Kong resembles the global scenario remains a speculation.

Our NoroBase has the advantages of free access, more comprehensive demographic and epidemiologically relevant information, coupled with complete viral genomic sequences. More functions such as alignment display and phylogenetic tree construction are under development. More hNoV strains other than the globally predominant GII.4 are being sequenced in our laboratory. Virus genotyping via

homology search against NoroBase may also benefit local epidemiologists and public health professionals in their studies of the epidemiologic pattern of hNoV infections in Hong Kong and in the identification of aetiologic agents in acute gastroenteritis outbreaks.

Acknowledgements

This study was supported by the Research Fund for the Control of Infectious Diseases, Food and Health Bureau, Hong Kong SAR Government (#03040482). We thank Dr WL Lim (Virology Division, Public Health Laboratory Centre, Department of Health, Hong Kong SAR Government) for providing human norovirus-containing faecal specimens.

References

1. Atmar RL, Estes MK. The epidemiologic and clinical importance of norovirus infection. Gastroenterol Clin North Am 2006;35:275-90.
2. Estes MK, Prasad BV, Atmar RL. Noroviruses everywhere: has something changed? Curr Opin Infect Dis 2006;19:467-74.
3. Ho EC, Cheng PK, Lau AW, Wong AH, Lim WW. Atypical norovirus epidemic in Hong Kong during summer of 2006 caused by a new genogroup II/4 variant. J Clin Microbiol 2007;45:2205-11.
4. Lau CS, Wong DA, Tong LK, et al. High rate and changing molecular epidemiology pattern of norovirus infections in sporadic cases and outbreaks of gastroenteritis in Hong Kong. J Med Virol 2004;73:113-7.
5. Ho EC, Cheng PK, Wong DA, Lau AW, Lim WW. Correlation of norovirus variants with epidemics of acute viral gastroenteritis in Hong Kong. J Med Virol 2006;78:1473-9.
6. Zheng DP, Ando T, Fankhauser RL, Beard RS, Glass RI, Monroe SS. Norovirus classification and proposed strain nomenclature. Virology 2006;346:312-23.
7. Wobus CE, Karst SM, Thackray LB, et al. Replication of Norovirus in cell culture reveals a tropism for dendritic cells and macrophages. PLoS Biol 2004;2:e432.
8. Jiang X, Espul C, Zhong WM, Cuello H, Matson DO. Characterization of a novel human calicivirus that may be a naturally occurring recombinant. Arch Virol 1999;144:2377-87.
9. Yee EL, Palacio H, Atmar RL, et al. Widespread outbreak of norovirus gastroenteritis among evacuees of Hurricane Katrina residing in a large "megashelter" in Houston, Texas: lessons learned for prevention. Clin Infect Dis 2007;44:1032-9.
10. Le Guyader FS, Bon F, DeMedici D, et al. Detection of multiple noroviruses associated with an international gastroenteritis outbreak linked to oyster consumption. J Clin Microbiol 2006;44:3878-82.
11. Parshionikar SU, Willian-True S, Fout GS, et al. Waterborne outbreak of gastroenteritis associated with a norovirus. Appl Environ Microbiol 2003;69:5263-8.
12. Soares CC, Santos N, Beard RS, et al. Norovirus detection and genotyping for children with gastroenteritis, Brazil. Emerg Infect Dis 2007;13:1244-6

KF To 杜家輝
WK Leung 梁偉強
KW Ng 吳國偉
JHM Tong 唐紅文
RWM Lung 龍偉明

Sequencing analysis of the 3′ region of the cagA gene in *Helicobacter pylori* isolated from Hong Kong Chinese patients

Key Messages

1. The cagA gene is suggested as a marker for *Helicobacter pylori* strains with enhanced virulence, and diversity in the 3' region of the gene may be associated with different biological activities of various strains.
2. We conducted the largest and most comprehensive series of sequence analyses of the *H pylori* cagA 3' region and documented sequence variants prevailing in Hong Kong.
3. Three subtypes (1, 2, and 3) of the cagA gene were classified based on the type and number of amino acid repeats in the 3' region. Although subtypes 2 and 3 (containing 4 copies of R1 EPIYA repeats) significantly correlated with aggressive disease, subtype 1 remained the most common strain in Hong Kong (96% of the samples).
4. 964-T/S/V variants carrying mutations in a residue adjacent to the tyrosine phosphorylation site were detected in benign lesions only. This observation should be confirmed in a larger sample, together with a functional analysis to elucidate the biological significance of these variants.

Hong Kong Med J 2010;16(Suppl 4):S8-12

The Chinese University of Hong Kong:
Department of Anatomical and Cellular Pathology
KF To, JHM Tong, RWM Lung
Department of Medicine and Therapeutics
WK Leung
Department of Surgery
KW Ng

RFCID project number: 03040552

Principal applicant and corresponding author:
Prof Ka Fai To
Department of Anatomical and Cellular Pathology, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, NT, Hong Kong SAR, China
Tel: (852) 2632 2361
Fax: (852) 2637 6274
Email: kfto@cuhk.edu.hk

Introduction

Helicobacter pylori (HP) is the aetiological agent causing chronic gastritis, peptic ulcer and gastric cancer. Strain-specific genetic diversity of HP is postulated to influence the clinical outcome of HP infection, but with considerable geographic differences. The cagA gene (cytotoxin-associated gene A) is considered a marker for a pathogenicity island of about 40 kbs.[1] The presence of cagA is associated with duodenal ulcer, gastric ulcer, gastric mucosal atrophy and gastric cancer. Variations in the number of repeats of the amino acid sequence R1 (EPIYA) have been identified in Japanese HP isolates.[2] The cagA genotype containing four or more R1 sequences is associated with atrophic gastritis and gastric cancer. Hong Kong has a high prevalence of HP infection and a high incidence of gastroduodenal diseases. About 90% of the HP strains are cagA positive. Since allelic variation in cagA exists and distinct HP subgenotypes may circulate in different regions, such differences may provide markers for differences in virulence among cagA-positive HP strains. Data on the sequence variations of HP cagA gene in Hong Kong inhabitants are scanty. A large-scale, comprehensive study is required to determine the genotypes of HP cagA gene in Hong Kong.

Methods

This study was conducted from January 2006 to January 2007. DNA of HP was extracted from both HP isolates and directly from gastric biopsy samples of subjects presenting with gastrointestinal symptoms (including dyspepsia and gastro-intestinal bleeding) and asymptomatic subjects for screening. Endoscopic features and biopsy findings were reviewed and samples were classified according to different HP-associated diseases. The presence of the HP genome and cagA were further confirmed by polymerase chain reaction (PCR). Direct DNA sequencing was performed to investigate sequencing variations of the cagA gene. The results were correlated with biopsy assessments and with different disease groups.

A total of 278 consecutive samples from subjects with HP-associated diseases were included. After reviewing relevant clinical data and biopsy features, the samples were classified into six disease categories: minimal-to-mild gastritis (n=17), moderate-to-severe gastritis (n=51), gastric ulcer (n=48), duodenal ulcer (n=88), severe atrophic gastritis with marked intestinal metaplasia (n=42), and gastric cancer (n=32).

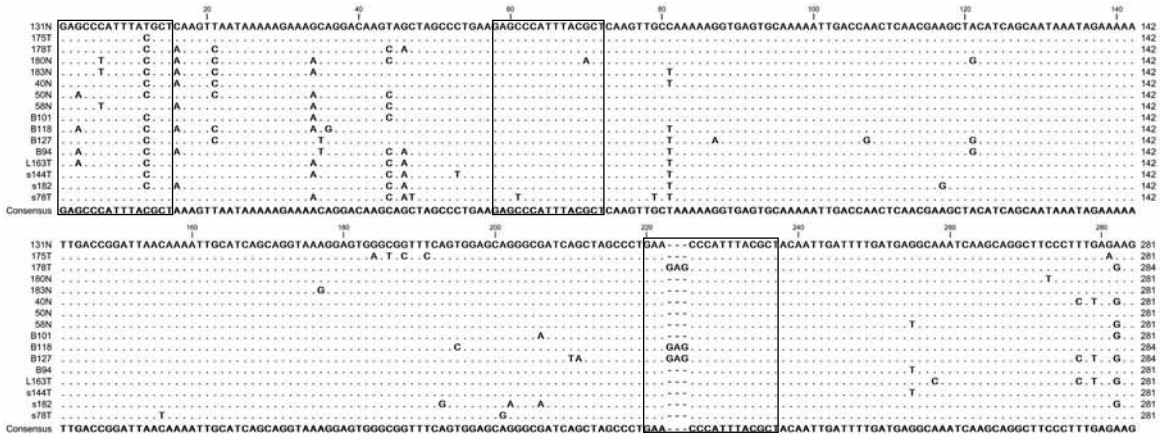
DNA was extracted from clinical samples using a High Pure PCR Template Preparation Kit (Roche) according to the manufacturer's instructions. Primer sets specific for the HP 16S rRNA gene were used for the detection of HP. The diversity of the cagA gene was examined by sequencing its 3' region.

Results

Of the 278 HP positive samples, 225 (81%) containing the cagA gene were classified as minimal-to-mild gastritis (n=17), moderate-to-severe gastritis

(a) Subtype 1 of HP cagA gene

Alignments of the nucleotide sequences



Alignments of the amino acid sequences

(b) Subtype 2 of HP cagA gene

Alignments of the nucleotide sequence

Alignments of the amino acid sequences

(c) Subtype 3 of HP cagA gene

Alignments of the nucleotide sequences

Alignments of the amino acid sequences

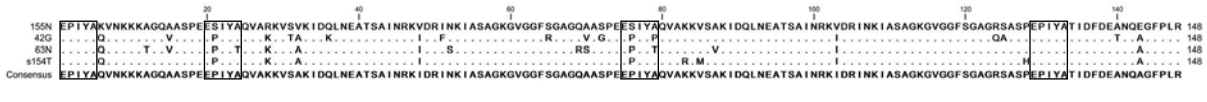


Fig 1. Alignments of the nucleotide and amino acid sequences of subtypes of the HP cagA gene

(n=33), gastric ulcer (n=41), duodenal ulcer (n=67), severe atrophic gastritis with marked intestinal metaplasia (n=37), and gastric cancer (n=30).

Sequencing analysis of the 225 HP cagA positive samples revealed that all the variants demonstrated characteristics of East Asian strains, which differed markedly from western strains.[3] The primary structure of the 3' region of the cagA gene was composed of a variable number of repeat sequences, including R1 (15-bp, 5 amino acids EPIYA), R2 (42-bp, 14 amino acids KVNKKKTGQVASPE) and R3 (147-bp, 49 amino acids, QVAKKVKAKIDQLNE ATSAINRKIDRINKIASAGKGVGGFSGAGQSASP). The structural organisation could be divided into three subtypes: subtype 1 (R1-R2-R1-R3-R1), subtype 2 (R1-R2-R1-R2-R1-R3-R1) and subtype 3 (R1-R2-R1-R3-R1-R3-R1). Alignments of the representative nucleotide and deduced amino acid sequences for each subtype are shown in Fig 1.

We investigated the diversity of the cagA gene among Hong Kong strains and the association between sequence diversity and clinical status. The distribution of cagA 3' region subtypes in different gastrointestinal diseases is shown in the Table. Subtype 1 strain with three copies of R1 (EPIYA) repeats was the most common (217/225, 96%). Both subtypes 2 and 3 contained four copies of R1 (EPIYA) repeats. Subtype 2 strain was present in four (2%) cases (one gastric cancer, two severe atrophic gastritis and one duodenal ulcer). Subtype 3 strain was present in another four (2%) cases (three gastric cancer and one duodenal ulcer).

The frequencies of the cagA genotype with four copies of R1 (EPIYA) repeats (subtypes 2 and 3) were significantly higher in isolates from patients with gastric cancers than with ulcers and gastritis (p=0.006, Fisher's exact test). As severe atrophic gastritis with marked intestinal metaplasia is regarded as pre-cancerous, we combined gastric cancer and severe atrophic gastritis as one disease group and compared corresponding cagA genotypes from persons with benign lesions (ulcer and gastritis). A significantly higher frequency of 4xR4 was also noted in the aggressive as opposed to the benign lesion group (odds ratio=7.7, 95% confidence interval=1.35-56.65, p=0.0096).

Biological diversities among different cagA proteins are caused by variations in the number and sequences of tyrosine phosphorylation sites in the molecule.[3] Thus, we tried to look for cagA strains with sequence variations adjacent to the tyrosine phosphorylation sites. CagA strains with amino acid changes from alanine to threonine/serine/valine at residue 964 (A964T/S/V) corresponding to East Asian strain F32-cagA (GenBank accession number AF202972) were identified. This was located at the eighth residue, immediately following phosphotyrosine (pY) within the last R1 (EPIYA) repeat. These strains were designated 964-T, 964-S and 964-V. The prototype was designated 964-A (Fig 2). The 964 variants were disproportionately found in patients with aggressive versus benign lesions. The 964-T variants were present in 12 (5%) cases (two gastritis, eight duodenal ulcers, and two gastric ulcers), whereas the 964-S variant was noted in one duodenal ulcer subject and 964-V variants were found in one patient with a duodenal ulcer and one with mild gastritis. No variants were detected in subjects with atrophic gastritis or gastric cancer. There was a trend towards a higher frequency of 964-T/S/V variants in persons with benign lesions (15/158 with ulcer and gastritis, 10%) rather than aggressive lesions (0/67 with atrophic gastritis and gastric caner) [P=0.007, Fisher's exact test].

Discussion

This study is the largest and most comprehensive series analysed for sequence variations at the 3' region of the HP cagA gene. A wide range of HP-associated gastrointestinal diseases from different subjects with various pathologies were studied. The cagA gene was detected in 225 (81%) out of the 278 HP positive strains. The frequency was consistent with reported frequencies from Korea and China, whereas in western countries, the percentage was much lower.

Sequencing analysis revealed three types of cagA gene (subtypes 1 to 3), which were identical to the variants from Zhejiang province, China.[4] Four subtypes of cagA isolated (types A to D) were reported from Japanese patients.[3] The Japanese types A and C were identical to our subtypes 1 and 3, respectively. However, we were not able to identify Japanese types B and D in our series. Similar findings were also reported elsewhere,[4] indicating distinct HP strains circulated in different geographic regions.

The HP cagA gene was heterogeneous with respect to its potential for undergoing tyrosine phosphorylation, SHP-2

Table. Distribution of the subtypes of the 3' region of the cagA gene in different gastroduodenal diseases

Subtype of 3' region of cagA gene	No. (%) of cagA positive cases						
	Gastric cancer (n=30)	Severe atrophic gastritis (n=37)	Gastric ulcer (n=41)	Duodenal ulcer (n=67)	Moderate-to-severe gastritis (n=33)	Minimal-to-mild gastritis (n=17)	Total (n=225)
Subtype 1	26 (87)	35 (95)	41 (100)	65 (97)	33 (100)	17 (100)	217 (96)
Subtype 2	1 (3)	2 (5)	0 (0)	1 (2)	0 (0)	0 (0)	4 (2)
Subtype 3	3 (10)	0 (0)	0 (0)	1 (2)	0 (0)	0 (0)	4 (2)
Subtypes 2 and 3	4 (13)	2 (5)	0 (0)	2 (3)	0 (0)	0 (0)	8 (4)

binding, and induction of cellular morphological changes. Such biological diversities among different cagA proteins were caused by variations in the number and sequences of tyrosine phosphorylation sites of the molecule. The R1 EPIYA repeat motifs in the 3' cagA region were potential targets of tyrosine phosphorylation. The cagA protein with more EPIYA repeats were expected to be more active biologically than those with a small number of repeats because they interacted more effectively with SHP-2 phosphatase and more severely perturb SHP-2-dependent signalling pathways. In keeping with this theory, our results demonstrated HP cagA strains carrying cagA protein with more repeat sequences (four versus three copies of R1 EPIYA) were associated with severe atrophic gastritis and gastric cancer.

Our results indicated that cagA strain subtypes 2 and 3 with four copies of R1 EPIYA repeat sequences correlated with aggressive lesions. Subtype 1 with three copies of R1 EPIYA repeats remained the most common pattern in Hong Kong, even in patients with severe atrophic gastritis and gastric cancer. Other virulence factors may contribute to different disease outcomes.

SH2 domains recognise phosphopeptide motifs composed of phosphotyrosine (pY) followed by several C-terminal residues. Alterations in the ligand-binding motifs might reduce the binding affinity to SHP-2. We identified two strain variants with amino acid changed from alanine to threonine/serine/valine at the residue 964 (A964T/S/V), the eighth residue immediately following phosphotyrosine (pY) within the last R1 (EPIYA) repeat. 964-T/S/V variants were detected in benign lesions only. Based on this observation, we proposed a previously uncharacterised model for the cagA-SHP-2 interaction. Change of the pY+8 residue in these variants might alter the SHP-2-binding affinity, thus reducing the ability to form a complex with SHP-2 compared with the complex-forming activity of wild-type cagA protein. Therefore, 964-T/S/V variants were less active biologically than wild-type counterparts. Interestingly, one

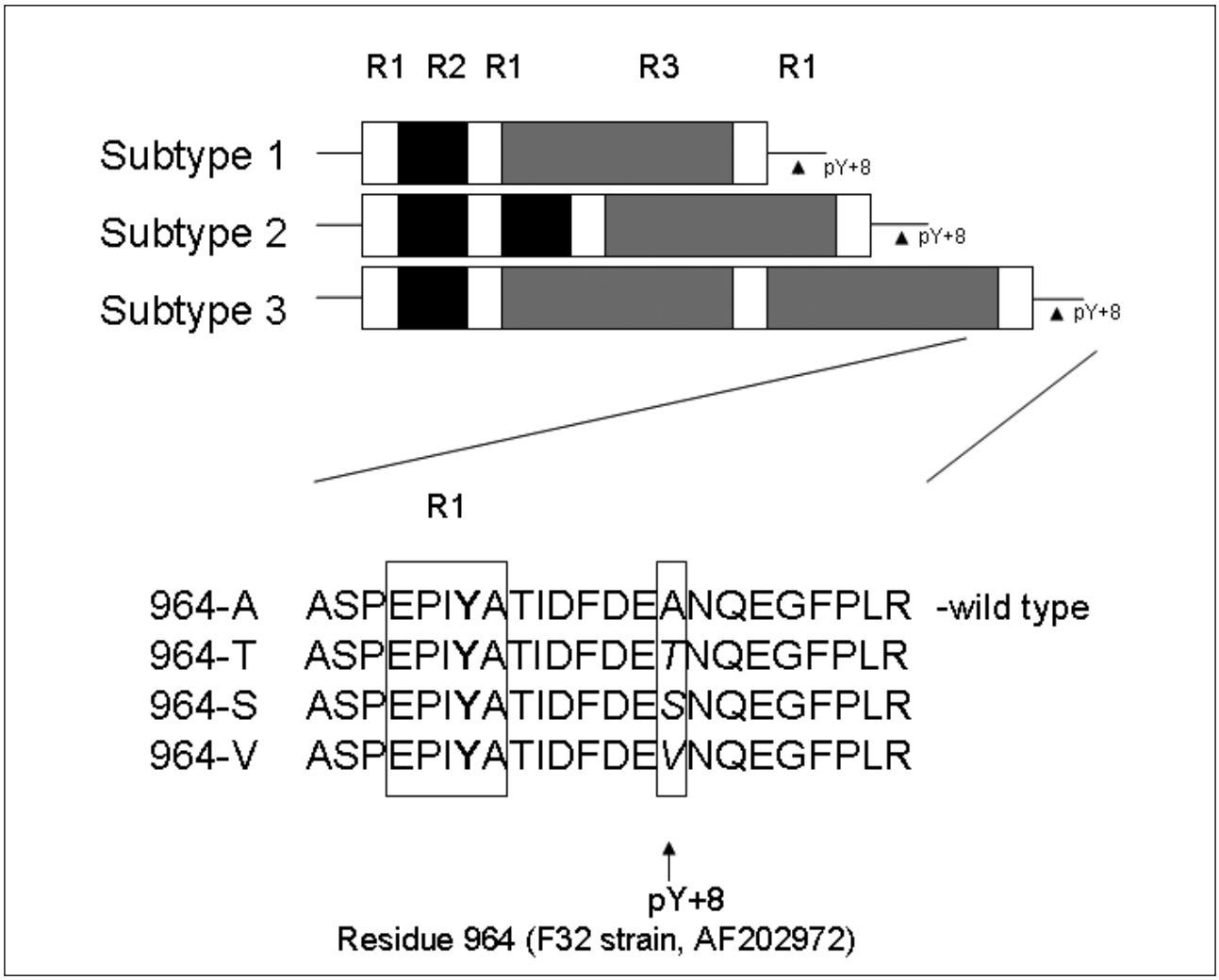


Fig 2. Schematic demonstration of pY+8 variant strains in three subtypes of the cagA protein
Arrowheads indicate the location of pY+8 residue immediately following the last R1 EPIYA repeats in each of the cagA subtypes. The phosphotyrosines are bolded in the sequence alignment of the three variants. The EPIYA repeats and pY+8 residues are boxed

of the 964-T variants belonged to the cagA subtype 3 with four copies of R1 EPIYA repeats. This strain was isolated from a patient with duodenal ulcer. cagA proteins with four copies of R1 EPIYA repeats were presumably more active biologically. The simultaneous possession of 964-T might reduce the SHP-2 binding affinity of this particular variant, rendering it less active. However, these observations need to be confirmed in studies with larger samples together with investigation of functional aspects, so as to further elucidate our hypothesis.

Acknowledgement

This study was supported by Research Fund for the Control of Infectious Diseases, Food and Health Bureau, Hong Kong SAR Government (#03040552).

References

1 Censini S, Lange C, Xiang Z, et al. cag, a pathogenicity island of *Helicobacter pylori*, encodes type I-specific and disease-associated virulence factors. Proc Natl Acad Sci U S A 1996;93:14648-53.

2 Azuma T, Yamakawa A, Yamazaki S, et al. Correlation between variation of the 3' region of the cagA gene in *Helicobacter pylori* and disease outcome in Japan. J Infect Dis 2002;186:1621-30.

3 Higashi H, Tsutsumi R, Fujita A, et al. Biological activity of the *Helicobacter pylori* virulence factor CagA is determined by variation in the tyrosine phosphorylation sites. Proc Natl Acad Sci U S A 2002;99:14428-33.

4 Tao R, Fang PC, Liu HY, Jiang YS, Chen J. A new subtype of 3' region of cagA gene in *Helicobacter pylori* strains isolated from Zhejiang Province in China. World J Gastroenterol 2004;10:3284-8.

SKP Lau 劉嘉珮
PCY Woo 胡釗逸
KY Yuen 袁國勇

Multi-locus sequence typing scheme for *Laribacter hongkongensis*, a novel bacterium associated with freshwater fish-borne gastroenteritis and traveller's diarrhoea

Key Messages

1. A multilocus sequence typing (MLST) system was developed for *Laribacter hongkongensis*, an emerging pathogen associated with fish-borne gastroenteritis and traveller's diarrhoea.
2. A website for *L hongkongensis* MLST at http://mlstdb.hku.hk:14206/MLST_index.html enables comparison of strains from different localities.
3. The same species of freshwater fish in a particular market is usually obtained from the same fish farm. The presence of multiple sequence types in *L hongkongensis* strains isolated from the same species of freshwater fish from the same market implies that multiple clones of *L hongkongensis* probably exist in the same fish farm in mainland China or Hong Kong.
4. The clustering of fish and human isolates into different groups observed previously using pulsed-field gel electrophoresis and the present MLST studies suggest that some clones of *L hongkongensis* are more virulent than others.

Hong Kong Med J 2010;16(Suppl 4):S13-4

Department of Microbiology, The University of Hong Kong
SKP Lau, PCY Woo, KY Yuen

RFCID project number: 04050112

Principal applicant:
Dr Susanna KP Lau
Department of Microbiology, The University of Hong Kong, University Pathology Building, Queen Mary Hospital, Pokfulam, Hong Kong SAR, China
Tel: (852) 2855 4892
Fax: (852) 2855 1241
Email: skplau@hkucc.hku.hk

Corresponding author:
Prof Kwok-yung Yuen
Department of Microbiology, The University of Hong Kong, University Pathology Building, Queen Mary Hospital, Pokfulam, Hong Kong SAR, China
Tel: (852) 2855 4892
Fax: (852) 2855 1241
Email: hkumicro@hkucc.hku.hk

Introduction

Laribacter hongkongensis was first discovered in Hong Kong in 2001 from the blood and empyema pus of a 54-year-old man with alcoholic cirrhosis and bacteraemic empyema thoracis.[1] It is a facultative anaerobic, motile, non-sporulating, urease-positive, Gram-negative, S-shaped bacillus. Genotypically, *L hongkongensis* belongs to the *Neisseriaceae* family of the β-subclass of Proteobacteria. It was subsequently discovered in the stool of six patients with community-acquired gastroenteritis. In a multi-centre prospective study using cefoperazone MacConkey agar as the selective medium, *L hongkongensis* was associated with community-acquired gastroenteritis.[2] Freshwater fish were confirmed to be a reservoir.[2]

L hongkongensis has been found in the intestines of healthy freshwater fish but not other studied animals that are commonly used for cooking in Hong Kong.[3] The bacterium was isolated from the guts of 24% out of 360 freshwater fish that were studied.[3,4] *L hongkongensis* has also been isolated from drinking water reservoirs in Hong Kong.[5] The presence of a heterogeneous population of *L hongkongensis* identified using pulsed-field gel electrophoresis (PFGE) of freshwater fish isolates[3] and the association of *L hongkongensis* gastroenteritis with fish consumption[2] suggests that freshwater fish are a likely major reservoir of the bacterium and the source of infections.

A highly reproducible and discriminative typing system is essential for better understanding of the epidemiology of *L hongkongensis*. We have used PFGE for typing *L hongkongensis*.[2,3,5] Due to experimental variations, PFGE patterns are difficult to compare among different laboratories. As multi-locus sequence typing (MLST) is highly reproducible and discriminative for bacteria, we developed such a typing system for *L hongkongensis*.

Methods

This study was conducted from December 2005 to December 2007. In order to have better understanding of the epidemiology of *L hongkongensis*, we developed an MLST system for *L hongkongensis*, using 146 human and fish isolates.[6] Seven housekeeping genes were selected for cross comparison: transcription termination facter Rho, aconitate hydratase, cell division protein, anthranilate synthase component I, ketol-acid reductoisomerase, thiamin biosynthesis protein ThiC , and enolase. The nucleotide sequences of seven housekeeping genes in all the *L hongkongensis* strains were aligned and compared.[6]

Regarding the 146 *L hongkongensis* isolates, 39 were from humans and 107 from fish. The fish isolates of were recovered from the gut of freshwater fish sampled from retail food markets in Hong Kong,[2,3] including 50 isolated from grass carp, 42 from bighead carp, 12 from mud carp and three from large-mouth bass.

Extracted DNA of the 146 isolates of *L hongkongensis* was used as the template for amplification of seven housekeeping genes.[6] The nucleotide sequences of the seven gene loci in all the *L hongkongensis* strains were aligned and compared with those of strain HLHK1 using Clustal W multiple alignment implemented in BioEdit (version 7.0.5.2). The ratio of non-synonymous to synonymous base substitutions (d_n/d_s) was calculated with START2 (Sequence Type Analysis and Recombinational Tests Version 2) [http://pubmlst.org/software/analysis/]. A dendrogram was constructed using the unweighted pair group method with the arithmetic mean using START2. Grouping of STs into lineages was performed with eBURST. The BURST algorithm first identifies mutually exclusive groups of related genotypes in the population (typically an MLST database), and attempts to identify the founding genotype (ST) of each group. The algorithm then predicts the descent from the predicted founding genotype to the other genotypes in the group. The linkage disequilibrium between alleles at the seven gene loci was measured using the standardised index of association (I^S_A) with LIAN 3.5 (http://www.mlst.net).

Results

Variations at the seven MLST loci

Among the 3068 bp of the seven loci, 332 polymorphic sites were observed in the 146 isolates of *L hongkongensis*.[6] Allelic profiles were assigned to the 146 isolates. The alleles defined for the MLST system were based on sequence lengths of between 362 bp (ilvC) and 504 bp (acnB).[6] No strong positive selective pressure was present at all seven loci.

Relatedness of L hongkongensis isolates

A total of 97 different STs were assigned to the 146 *L hongkongensis* isolates.[6] Their discriminatory power was 0.9861. eBURST grouped the isolates into 12 lineages, with 14 STs in group 1, 12 in group 2, seven in group 3, three in groups 4 to 6, and two in groups 7 to 12, whereas 43 STs did not belong to any of the 12 groups.[6] A website for *L hongkongensis* MLST at http://mlstdb.hku.hk:14206/MLST_index.html was set up.

No relationships were observed among the *L hongkongensis* isolates with respect to their year of isolation, location of the respective hospitals, age and gender of the patients, the presence of plasmids in the patient strains, the species of fish, and the locations of markets where the fish were purchased.

Discussion

The clustering of fish and human isolates into different groups observed in both the previous PFGE and the present MLST studies suggested that some clones of *L hongkongensis* were more virulent than others. Differential gene expression experiments using strains of human- and fish-derived isolates may reveal *L hongkongensis* virulence factors.

A heterogeneous population of *L hongkongensis* existed in the same ecosystem. *L hongkongensis* isolated from neither the same species of fish nor the same fish market was clustered together. Over 80% of freshwater fish consumed in Hong Kong are imported from fish farms in mainland China, whereas the remaining 20% are locally reared in fish farms in rural areas of Hong Kong. The same species of freshwater fish in a particular market was usually obtained from the same fish farm, but multiple STs were present in *L hongkongensis* strains isolated from the same species of freshwater fish purchased from the same market. This implies that multiple clones of *L hongkongensis* probably exist in the same fish farm in mainland China or Hong Kong.

Acknowledgement

This study was supported by the Research Fund for the Control of Infectious Diseases, Food and Health Bureau, Hong Kong SAR Government (#04050112).

References

1. Yuen KY, Woo PC, Teng JL, Leung KW, Wong MK, Lau SK. *Laribacter hongkongensis* gen. nov., sp. nov., a novel Gram-negative bacterium isolated from a cirrhotic patient with bacteremia and empyema. J Clin Microbiol 2001;39:4227-32.
2. Woo PC, Lau SK, Teng JL, et al. Association of *Laribacter hongkongensis* in community-acquired human gastroenteritis with travel and eating fish: a multicentre case-control study. Lancet 2004;363:1941-7.
3. Teng JL, Woo PC, Ma SS, et al. Ecoepidemiology of *Laribacter hongkongensis*, a novel bacterium associated with gastroenteritis. J Clin Microbiol 2005;43:919-22.
4. Lau SK, Woo PC, Fan RY, Lee RC, Teng JL, Yuen KY. Seasonal and tissue distribution of *Laribacter hongkongensis*, a novel bacterium associated with gastroenteritis, in retail freshwater fish in Hong Kong. Int J Food Microbiol 2007;113:62-6.
5. Lau SK, Woo PC, Fan RY, et al. Isolation of *Laribacter hongkongensis*, a novel bacterium associated with gastroenteritis, from drinking water reservoirs in Hong Kong. J Appl Microbiol 2007;103:507-15.
6. Woo PC, Teng JL, Tsang AK, et al. Development of a multi-locus sequence typing scheme for *Laribacter hongkongensis*, a novel bacterium associated with freshwater fish-borne gastroenteritis and traveler's diarrhea. BMC Microbiol 2009;9:21.

PCY Woo 胡釗逸
KY Yuen 袁國勇
SKP Lau 劉嘉珮
H Tse 謝勰文

Random exploration of the *Laribacte hongkongensis* genome

Key Messages

1. Random exploration of the *Laribacter hongkongensis* genome was performed. A total of 1957 random sequence tags were sequenced and analysed, which represents about 26% of the *L hongkongensis* genome.
2. Random exploration of the genome can be completed at a fraction of the complete genome costing. The limited partial sequence obtained in this study revealed ample evidence about the genetic composition and potential virulence factors of *L hongkongensis*.
3. Different classes of potential virulence factors were identified. Among these, a complete urease cassette, which is a major virulence factor in another gastrointestinal pathogen, *Helicobacter pylori*, was observed.
4. Random exploration also provided a foundation for the complete genome sequencing of *L hongkongensis*.

Hong Kong Med J 2010;16(Suppl 4):S15-7

Department of Microbiology, The University of Hong Kong
PCY Woo, KY Yuen, SKP Lau, H Tse

RFCID project number: 05050012

Principal applicant:
Prof Patrick CY Woo
Department of Microbiology, The University of Hong Kong, University Pathology Building, Queen Mary Hospital, Pokfulam, Hong Kong SAR, China
Tel: (852) 2855 4892
Fax: (852) 2855 1241
Email: pcywoo@hkucc.hku.hk

Corresponding author:
Prof Kwok-yung Yuen
Department of Microbiology, The University of Hong Kong, University Pathology Building, Queen Mary Hospital, Pokfulam, Hong Kong SAR, China
Tel: (852) 2855 4892
Fax: (852) 2855 1241
Email: hkumicro@hkucc.hku.hk

Introduction

Laribacter hongkongensis, a novel genus and species, was first discovered in Hong Kong in 2001.[1] Phenotypically, it is a facultative anaerobic, motile, non-sporulating, urease-positive, Gram-negative, S-shaped bacillus. Phylogenetic analysis using 16S rRNA gene sequences revealed that *L hongkongensis* belongs to the *Neisseriaceae* family of the β-subclass of Proteobacteria. During a period of 2 months, *L hongkongensis* was discovered, on charcoal cefoperazone deoxycholate agar, in three of our patients with community-acquired gastroenteritis. A similar finding was also observed in three other patients in Switzerland.[2] In a multi-centre prospective study using cefoperazone MacConkey agar as the selective medium,[3] we confirmed that *L hongkongensis* is associated with community-acquired gastroenteritis and traveller's diarrhoea.[4] Furthermore, freshwater fish were also confirmed to be a reservoir of *L hongkongensis*.[4,5] *L hongkongensis* is likely to be globally distributed, as travel histories from patients suggested that it was present in at least four continents, including Asia, Europe, Africa and Central America.[2,4] *L hongkongensis* has also been reported from another coastal province in mainland China. Although the causative role of *L hongkongensis* in gastroenteritis is yet to be established, these data provide strong evidence that the bacterium is a potential diarrhoeal pathogen that warrants further investigation.

Sequencing the complete genome of microorganisms has revolutionised the study of microbiology and infectious disease. However, due to the relative high cost of sequencing a complete genome, sample sequences of bacterial genomes have been used for characterisation of the microorganism. Recently, we published a sample sequence of *Penicillium marneffei*, which has facilitated further molecular research and provided the foundation for complete genome sequencing of this dimorphic fungus. In this study, random exploration of the *L hongkongensis* genome was performed, and 1957 random sequence tags (RSTs) were sequenced and analysed, which represents about 26% of the *L hongkongensis* genome. This study has laid down the foundation of complete genome sequencing of this bacterium.[6]

Methods

This study was conducted from July 2006 to December 2007. In order to have better understanding of the biology and putative virulence mechanisms of *L hongkongensis*, we sequenced and annotated 1957 RSTs, which represents about 26% of the *L hongkongensis* genome. The DNA sequences of the resultant contigs and singlets were analysed using the Basic Local Alignment Search Tool (BLAST) from the National Center for Biotechnology Information server at the National Library of Medicine (http://www.ncbi.nlm.nih.gov). Special attention was paid to comparison with genomes and other known virulence genes responsible for adhesion, survival, diarrheogenesis in *Chromobacterium violaceum* (a bacterium phylogenetically closely related to *L hongkongensis* with its complete genome sequence available) and other diarrhoeal pathogens.

Single sequencing reads, with an average length of about 500 bp, were obtained for one end of 2000 clones, using the T3 vector primer of pBK-CMV.

Using the sequence information of the sample sequences, additional

polymerase chain reaction (PCR) primers (Life Technologies Corp, CA, USA) were designed to amplify the gaps of the urease cassette of *L hongkongensis*.

DNA sequencing was performed with an ABI 3700 automated sequencer according to the manufacturer's instructions (Life Technologies Corp, CA, USA). To remove low-quality traces, the raw sequence reads were processed by *ad hoc* scripts. After filtering, the entire sets of RSTs were assembled into contigs using the Phred/Phrap/Consed software. The DNA sequences of the resultant contigs and singlets were analysed by BLAST search. The searches were performed at both the protein and DNA levels. All results were inspected and interpreted manually.

Phylogenetic tree construction of the predicted protein products of the urease cassette was performed using ClustalX (version 1.81) and the neighbour-joining method with GrowTree (Genetics Computer Group, Inc, WI, USA).

Results

General features of the random sequence tags

A random analysis of 1957 RSTs representing about 26% of the genome of *L hongkongensis* was performed. Clusters of Orthologous Groups (COG) classification showed that 6.2% of the genes had putative functions related to translation, 5.1% to transcription, 4.9% to replication, recombination and repair, 1.2% to cell cycle control, mitosis and meiosis, 1.3% to defence mechanisms, 5% to signal transduction mechanisms, 6.2% to cell wall and cell membrane biogenesis, 2.2% to cell motility, 1.8% to intracellular trafficking and secretion, 4.1% to posttranslational modification, protein turnover and chaperones, 6.4% to energy production and conversion, 2.7% to carbohydrate transport and metabolism, 8.3% to amino acid transport and metabolism, 2.1% to nucleotide transport and metabolism, 4.1% to coenzyme transport and metabolism, 3.1% to lipid transport and metabolism, 5.1% to inorganic ion transport and metabolism, 1.5% to secondary metabolites biosynthesis, transport and catabolism, 10.5% to general function prediction only, 7.5% to unknown functions, and 10.8% were not in the COGs. Phylogenetically, a large proportion of the genes were homologous to those of other members of the *Neisseriaceae* family of beta-proteobacteria, most commonly, *C violaceium*, the most closely related bacterium whose complete genome is available. Similar to *C violaceum*, *L hongkongensis* possessed a higher proportion of genes with putative functions related to signal transduction mechanisms (COG:T).

Potential virulence factors

Different classes of potential virulence factors were observed. They included genes that putatively encode urease, haemolysins, RTX toxin, phospholipase and adhesin. They also included a diverse array of protein secretion, iron uptake, efflux pump and lipopolysaccharide production systems.

Using the information of the sample sequences, additional PCR primers were designed to amplify the gaps of the urease cassette of *L hongkongensis*. The complete urease cassette, occupying a 7556 bp region, includes eight ORFs which encode three urease structural proteins (UreA, UreB and UreC) and five accessory proteins (UreE, UreF, UreG, UreD and UreI). Similar to the urease of other bacteria, the urease of *L hongkongensis* is presumably a nickel-containing enzyme. The histidine residues at the carboxyl terminal of UreE are supposed to bind to the nickel ions that are transported into *L hongkongensis* through a nickel transporter, and donate the nickel ions to UreC during urease activation. Phylogenetically, all seven genes in the urease cassette are closely related to the corresponding homologues in *Brucella* species, *Yersinia* species and *Photorhabdus luminescens* subsp *laumondii*.

Discussion

The complete sequencing of bacterial genomes has revolutionised microbiology. However, the current high cost of completely sequencing genomes has limited application of such technology to important pathogens and commercially important bacteria. The majority of this cost is incurred owing to the labour-intensive methods, which must still be used to close gaps covering the last few percent of the genome and to reduce error rates to below 0.1%. In contrast, a partial sequence of a bacterial genome can be obtained at low cost, which makes it possible to consider sample sequencing of multiple genomes within a species, genus, or family. When a completely sequenced genome and a closely related sample-sequenced genome are compared, it is possible to identify sequences in the sampled genome that are already characterised in the completely sequenced genome. The vast GenBank database can be considered a huge collection of sample sequences for these purposes.

In addition to the cost advantage of partial genome sequencing, the limited partial sequences obtained in this study revealed ample evidence about the genetic composition and the potential virulence factors of *L hongkongensis*. With about 26% of the genome sequenced, the predicted coding ORFs can be classified into all COG functional categories, except for genes of RNA processing and modification. As for *C violaceum*, analyses indicated that a substantial proportion of the genes in the *L hongkongensis* genome had putative functions related to signal transduction (COG T), which are related to bacterial survival under diverse environmental conditions. This is in line with the lifestyles of the bacteria. *C violaceum* is a highly versatile, soil- and water-borne, free-living bacterium and therefore requires the highest proportion of genes devoted to this COG group. *Neisseria meningitidis* and *Neisseria gonorrhoeae*

are strictly aerobic bacteria with humans as the only host, and therefore they require the lowest proportion of genes. *L hongkongensis* can survive in human and freshwater fish and therefore require an intermediate proportion of genes with putative functions related to signal transduction.

Different classes of potential virulence factors were identified. Among these, a complete urease cassette, which is a major virulence factor in another gastrointestinal pathogen, *Helicobacter pylori*, was observed in the *L hongkongensis* genome. Urease is able to hydrolyse the limited amount of urea available in the stomach to generate carbon dioxide and ammonia, which increases the pH. Therefore, the complete urease cassette observed in the *L hongkongensis* genome probably relates to its survival in the gastrointestinal tract, as the bacterium has to pass through the highly acidic environment of the stomach before reaching the intestine. Bioinformatic analysis showed that all seven genes in the urease cassette are closely related to the corresponding homologues in *Brucella* species (α-proteobacteria), *Yersinia* species (γ-proteobacteria) and *P luminescens* subsp *laumondii* (γ-proteobacteria), instead of those in other members of β-proteobacteria. This indicates that *L hongkongensis* has probably acquired the genes through horizontal gene transfer after evolution into a distinct species.

Acknowledgement

This study was supported by the Research Fund for the Control of Infectious Diseases, Food and Health Bureau, Hong Kong SAR Government (#05050012).

References

1. Yuen KY, Woo PC, Teng JL, Leung KW, Wong MK, Lau SK. *Laribacter hongkongensis* gen. nov., sp. nov., a novel Gram-negative bacterium isolated from a cirrhotic patient with bacteremia and empyema. J Clin Microbiol 2001;39:4227-32.
2. Woo PC, Kuhnert P, Burnens AP, et al. *Laribacter hongkongensis:* a potential cause of infectious diarrhea. Diagn Microbiol Infect Dis 2003;47:551-6.
3. Lau SK, Woo PC, Hui WT, et al. Use of cefoperazone MacConkey agar for selective isolation of *Laribacter hongkongensis*. J Clin Microbiol 2003;41:4839-41.
4. Woo PC, Lau SK, Teng JL, et al. Association of *Laribacter hongkongensis* in community-acquired human gastroenteritis with travel and with eating fish: a multicentre case-control study. Lancet 2004;363:1941-7.
5. Teng JL, Woo PC, Ma SS, et al. Ecoepidemiology of *Laribacter hongkongensis*, a novel bacterium associated with gastroenteritis. J Clin Microbiol 2005;43:919-22.
6. Woo PC, Lau SK, Tse H, et al. The complete genome and proteome of *Laribacter hongkongensis* reveal potential mechanisms for adaptations to different temperatures and habitats. PLoS Genet 2009;5:e1000416.

WK Leung 梁偉強
PKS Chan 陳基湘
NLS Lee 李禮舜
JJY Sung 沈祖堯

Development of an *in vitro* cell culture model for human noroviruses and its clinical application

Key Messages

1. Despite the encouraging results on murine norovirus 1, there is no evidence for the successful cultivation of human norovirus (hNoV) in cell lines of haematopoietic lineage. Further effort to develop an *in vitro* cell culture system for hNoV is needed.
2. Using *ex vivo* cultures, we found that hNoV displays a previously unreported marked tropism to human duodenal glandular epithelial cells. Future research to evaluate the permissiveness of human cell lines of this type for hNoV infection is warranted.

Introduction

Human norovirus (hNoV), a member of the family *Caliciviridae* in the genus *Norovirus*, is the leading cause of acute non-bacterial gastroenteritis worldwide, and affects all age groups in both developed and developing countries. An overwhelming majority of sporadic cases as well as outbreaks were attributed to strains classified into genogroup II genotype 4 (GII.4). In the past decades, research on the development of a culture system or a small animal model for hNoV replication has made little progress. The failure of culturing hNoV has severely hampered the study of the pathogenesis of this important enteric pathogen and corresponding antiviral strategies. Attempts to cultivate hNoV using a broad panel of cell lines have been unsuccessful.[1] It has been reported that hNoV can infect differentiated human embryonic small intestinal cell line (INT 407) when grown as a three-dimensional culture,[2] but whether productive viral replication occurs remains unclear.[3]

Murine norovirus 1 (MNV-1) has recently been demonstrated to infect and replicate in murine macrophage cell lines and cultured primary dendritic cells and macrophages from STAT1$^{-/-}$ mice.[4] This unexpected virus tropism to cells of haematopoietic lineage has provided new insights into the possible development of an *in vitro* culture system for hNoV. We therefore set out to establish an *in vitro* cell culture model for hNoV based on the hypothesis that human and murine NoVs share similar tropism to cells of haematopoietic lineage. We summarise our efforts to cultivate hNoV in these cell types *in vitro* and detail our exploration of alternative approaches.

Study design and methods

This study was conducted from January 2006 to December 2007. We investigated the permissiveness of five haematopoietic cell lines (KG-1, THP-1, J774A.1, RAW264.7, and WBC264-9C) of human and murine origin for a panel of diverse hNoV strains that covered two genogroups and seven genotypes. We assessed successful viral infection and productive RNA replication by light microscopy and reverse transcription-quantitative polymerase chain reaction (RT-qPCR) targeting the junction between ORF1 and ORF2 of the viral genome. We also evaluated whether *STAT1* mRNA silencing makes the cell lines permissive for or better able to support hNoV replication. Furthermore, we evaluated whether *ex vivo* culture of human adult duodenal tissues can be used for the study of hNoV. We assessed successful viral infection and productive replication using RT-qPCR, *in situ* hybridisation, and immunohistochemical staining.

Results

Non-permissiveness of different cell lines of haematopoietic lineage for hNoV

Apparent CPE was observed in cell lines RAW264.7 and WBC264-9C when compared with mock inoculated cells and with cells inoculated with heat-inactivated virus (data not shown). However, only cell enlargement (a morphological change) was noted (data not shown). There was no evidence of cell lysis or foci formation. Using RT-qPCR, in all the cell lines tested, we

Hong Kong Med J 2010;16(Suppl 4):S18-21

The Chinese University of Hong Kong:
Department of Medicine and Therapeutics
WK Leung, NLS Lee, JJY Sung
Department of Microbiology
PKS Chan

RFCID project number: 04050382

Principal applicant and corresponding author:
Dr Wai Keung Leung
Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Shatin, NT, Hong Kong SAR, China
Tel: (852) 2632 3140
Fax: (852) 2637 3852
Email: dr_wkleung@alumni.cuhk.net

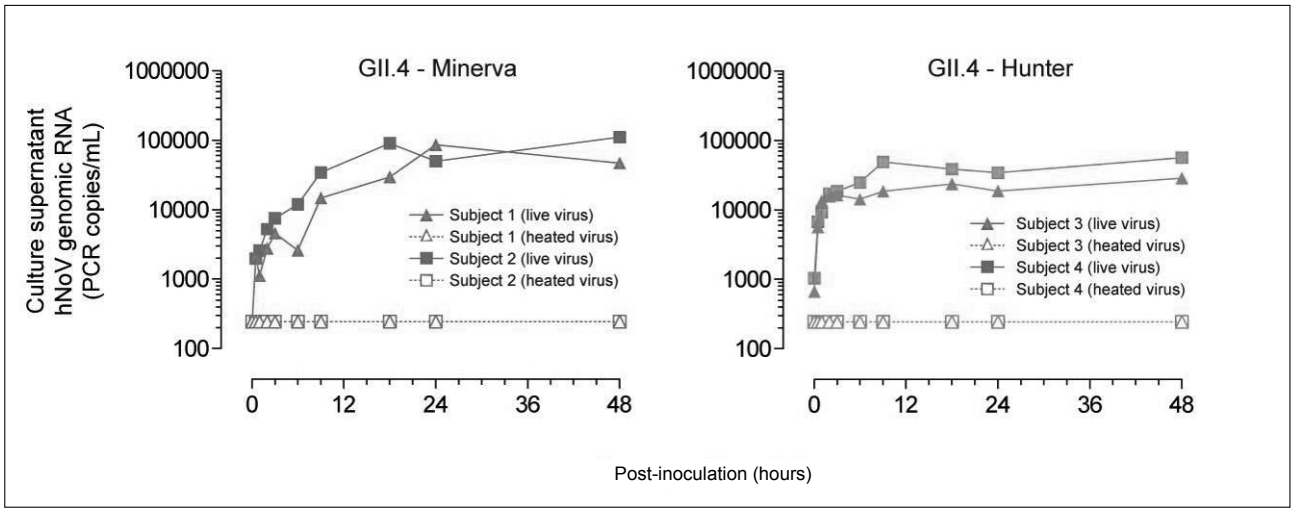


Fig 1. RNA growth kinetics of human norovirus hNoV) in *ex vivo*–inoculated adult human duodenal tissues
Viral RNA growth curves of two recent epidemic II.4 strains (Minerva and Hunter) are shown. Viral genomic RNA levels in cell-free culture supernatants are measured by RT-qPCR targeting the ORF1/2 junction

observed no substantial increase (>5 fold) in viral genomic RNA level up to 7 days post-inoculation (data not shown). Moreover, plaque assays were all negative (data not shown). Similar results were observed in all the five cell lines when STAT1 mRNA was silenced even after achieving substantial reductions of the mRNA and protein levels of STAT1, as confirmed by RT-qPCR and Western blotting up to 3 days after siRNA silencing (data not shown). Collectively, our data did not support the proposition that the five haematopoietic cell lines investigated were permissive for hNoV.

Infection and productive replication of hNoV in ex vivo culture using adult human duodenal tissues

We explored an alternative approach to study hNoV replication. We evaluated whether *ex vivo* culture can be used for the study of hNoV by inoculating freshly collected human duodenal tissues obtained from adult dyspeptic volunteers using two recent epidemic hNoV GII.4 strains (Minerva and Hunter). Aliquots of *ex vivo* culture supernatant were collected at 0, 0.5, 1, 2, 3, 6, 9, 18, 24, and 48 hours post-inoculation and clarified by centrifugation. We found that viral genomic RNA levels in cell-free culture supernatants increased over time as measured by RTqPCR (Fig 1).

To determine whether the productive hNoV RNA replication in *ex vivo* culture was accompanied by viral proteins expression, we performed immunohistochemical staining against structural VP1 as well as non-structural viral material. We detected intracellular cytoplasmic expressions of both VP1 (Fig 2a) and viral protease (Fig 2b) 48 hours post-inoculation. Detection of newly synthesised viral protease further supports the possibility that viral replication was occurring *ex vivo*.

We then performed *in situ* hybridisation to detect and locate intracellular hNoV RNA on tissues that were formalin-fixed and paraffin-embedded 48 hours post-inoculation. Digoxigenin-labelled antisense RNA probe that hybridises to and detects positive-stranded viral genomic RNA by targeting the ORF1/2 junction gave positive signals in both the cytoplasm and nuclei of infected cells (Fig 3a). This corroborates our RT-qPCR findings that hNoV can infect human duodenal tissues in *ex vivo* culture. Notably, positive signals were observed predominantly in glandular epithelial cells and were less frequently noted in cells on luminal epithelial surfaces (Fig 3a). The results showed that hNoV had a marked tropism for glandular epithelial cells of the human duodenum. This cell tropism was confirmed by comparable immunohistochemical staining patterns of structural VP1, and non-structural viral protease produced by *in situ* hybridisation (Fig 2). We also observed positive staining using the sense RNA probe that hybridises to and detects negative-stranded viral antigenomic RNA (Fig 3b). This indicates that hNoV was actively replicating in *ex vivo* cultures, through the production of negative-stranded RNA.

Limited viral RNA replication in a glandular epithelial cell line

We then selected and investigated the permissiveness of a glandular epithelial cell line known as human intestinal epithelial cell line 6 (HIEC-6) for a GII.4 strain of hNoV (Hunter) as described above. Limited productive RNA replication as evident by an up to 2-$\log_{10}$ increase in viral genomic RNA level in cell-free culture supernatant was detected 48 hours post-inoculation (data not shown). However, no observable CPE was noted in this cell line.

Discussion

The absence of a culture system for hNoV propagation and the non-permissiveness of replicon-based reverse genetics systems for this infection have severely hampered study of

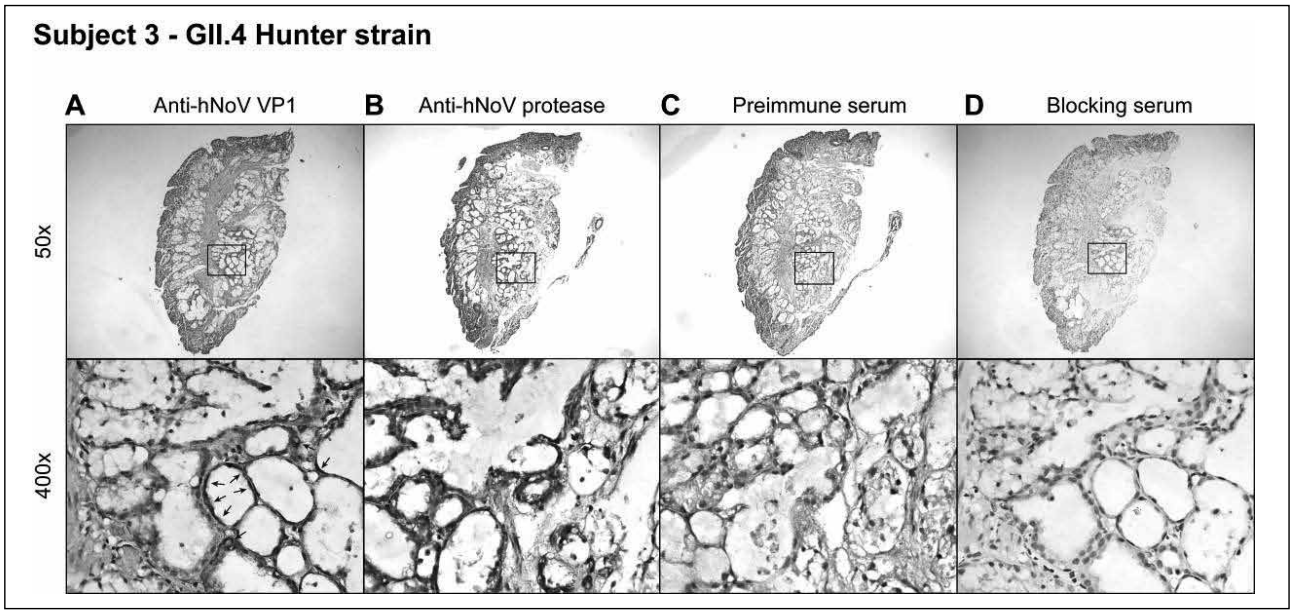


Fig 2. Immunoperoxidase staining for human norovirus hNoV) structural VP1 and newly synthesised non-structural viral protease in *ex vivo*–inoculated adult human duodenal tissues

(a and b) Formalin-fixed paraffin-embedded sections (48 hours post-inoculation) are stained with either mouse monoclonal anti-hNoV VP1 antibody, (a) NS14 or (b) rabbit polyclonal anti-hNoV protease antiserum, followed by DAB colour development. Positive signals (dark patches or dots) for VP1 (arrows) and viral protease are observed in the cytoplasm of infected cells resembling glandular epithelial cells. Sections are counterstained with Harris haematoxylin. (c) Preimmune serum of anti-hNoV protease antiserum and (d) blocking goat serum in replace of primary antibody are used as antibody specificity controls. Sections from volunteer subject 3 infected with Hunter strain are shown

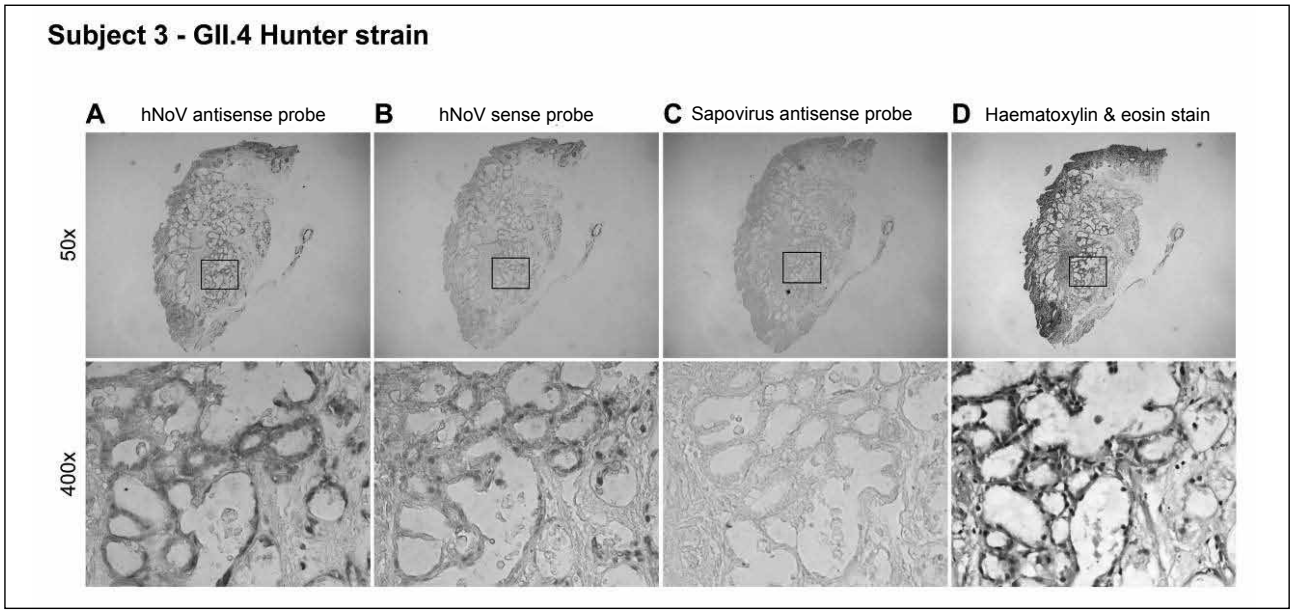


Fig 3. *In situ* hybridisation for human norovirus hNoV) genomic and antigenomic RNA in *ex vivo*–inoculated adult human duodenal tissues

(a and b) Formalin-fixed paraffin-embedded sections (48 hours post-inoculation) are hybridised with (a) digoxigenin-labelled antisense or (b) sense hNoV RNA probe targeting the ORF1/2 junction, followed by NBT/BCIP colour development. Positive signals (dark patches or dots) are noted in both the cytoplasm and nuclei of infected cells resembling glandular epithelial cells using antisense probe that hybridises to and detects (a) positive-stranded viral genomic RNA but are only noted in nuclei of infected cells using sense probe that hybridises to and detects (b) negative-stranded viral antigenomic RNA. (c) Sections are hybridised with unrelated digoxigenin-labelled antisense RNA probe targeting capsid gene sequence of human sapovirus. No observable signals are detected, confirming the positive signals in panels A and B are hNoV-specific. (d) Sections are stained with Harris haematoxylin and eosin

its pathogenesis and putative antiviral strategies.

In this study, five cell lines of haematopoietic lineage were investigated, after they were selected based on previous work that showed that MNV-1 can infect and replicate in them.[4] We hypothesised that human and murine NoVs share similar tropism to cells of haematopoietic lineage. Although apparent CPE was observed in cell lines RAW264.7 and WBC264-9C, no productive RNA replication was detected using RT-qPCR. Subsequent investigation of the apparent CPE showed no plaque formation. Since haematopoietic cells isolated from STAT1$^{-/-}$ mice had been shown to better support MNV-1 replication,[4] we evaluated whether silencing *STAT1* mRNA can render the cell lines permissive for hNoV. However, findings observed were similar to those with their non-silenced counterparts. We cannot exclude the possibility that siRNA treatment itself may pre-activate STAT1-independent anti-viral interferon response that in turn may inhibit hNoV replication in these silenced cell lines. The non-permissiveness of their non-silenced counterparts have been shown by other groups to readily support murine NoV replication, which casts serious doubts in regard to further investigation of these haematopoietic cell lines for hNoV cultivation. Collectively, our data do not support the notion that haematopoietic cell lines can support hNoV replication *in vitro*.

We then explored the use of an alternative approach to study hNoV. *Ex vivo* culture using freshly collected human tissues usually mimics the primary site of viral replication *in vivo*. This method has been successfully applied to studies of infection and replication of a diverse range of clinically important viruses such as the highly pathogenic avian influenza (H5N1) virus.[5] In this regard, we performed *ex vivo* culture of two recent epidemic GII.4 strains using freshly collected adult human duodenal tissues obtained from adult dyspeptic volunteers. Our findings illustrate that *ex vivo* culture can support key stages of complete hNoV replication, ranging from virus adsorption and internalisation to viral RNA replication and protein synthesis. Notably, *in situ* hybridisation of viral RNA and immunohistochemical staining of VP1 and newly synthesised viral protease revealed that hNoV displays a marked tropism for glandular epithelial cells. This previously unreported cell tropism suggests that human cell lines derived from or resembling intestinal glandular epithelial cells may support hNoV replication and merit investigation.

In this regard, we selected the glandular epithelial cell line HIEC-6 based on our ex vivo culture findings and evaluated its permissiveness for hNoV. HIEC-6 was human intestinal epithelial cells isolated from foetal ileal tissue. This was shown to express intestinal crypt cell marker antigen MIM-1/39.[6] Although CPE was not observed upon hNoV inoculation, limited viral RNA replication of a hNoV GII.4 strain (Hunter) was noted in this cell line. This suggests that cell lines of similar types are promising candidates worthy of further investigation. We are screening other glandular epithelial cell lines for their ability to support hNoV replication and, more importantly induce cytopathic effects that are crucial to the development of a plaque assay to quantify viable virus.

Acknowledgements

This study was supported by the Research Fund for the Control of Infectious Diseases, Food and Health Bureau, Hong Kong SAR Government (#04050382). We thank Dr Beaulieu of University of Sherbrooke, Canada for providing human intestinal epithelial cell line 6.

References

1. Duizer E, Schwab KJ, Neill FH, Atmar RL, Koopmans MP, Estes MK. Laboratory efforts to cultivate noroviruses. J Gen Virol 2004;85:79-87.
2. Straub TM, Honer zu Bentrup K, Orosz-Coghlan P, et al. In vitro cell culture infectivity assay for human noroviruses. Emerg Infect Dis 2007;13:396-403.
3. Chan MC, Wong YP, Leung WK. Cell culture assay for human noroviruses. Emerg Infect Dis 2007;13:1117.
4. Wobus CE, Karst SM, Thackray LB, et al. Replication of Norovirus in cell culture reveals a tropism for dendritic cells and macrophages. PLoS Biol 2004;2:e432.
5. Nicholls JM, Chan MC, Chan WY, et al. Tropism of avian influenza A (H5N1) in the upper and lower respiratory tract. Nat Med 2007;13:147-9.
6. Perreault N, Jean-Francois B. Use of the dissociating enzyme thermolysin to generate viable human normal intestinal epithelial cell cultures. Exp Cell Res 1996;224:354-64.

ZD Ye 葉仲德
CK Wong 黃仲佳
P Li 李　蓬
Y Xie 謝　雍

The role of SARS-CoV protein, ORF-6, in he induction of host cell death

Key Messages

1. Overexpression of ORF-6 induced apoptosis.
2. Caspase-3 inhibitor and JNK inhibitor blocked ORF-6 induced apoptosis.
3. The protein level of ER chaperon protein, GRP94, was up-regulated when ORF-6 was overexpressed.
4. ORF-6 induced apoptosis via caspase-3–mediated, ER stress and JNK dependent pathways.

Introduction

In 2003, Hong Kong experienced an outbreak of severe acute respiratory syndrome (SARS) and incurred huge economic and social losses. Immunohistochemistry and *in situ* hybridisation of organs from deceased SARS patients revealed that the virus was not only in the lungs and intestines, but also in the liver, distal convoluted renal tubules, sweat glands, parathyroid glands, pituitary gland, pancreas, adrenal glands, and the cerebrum. The aetiological agent for this disease was a novel member of the coronavirus (SARS-CoV) family, with limited sequence homology to other coronaviruses.[1-8] Coronaviruses are enveloped, plus-stranded RNA viruses that lead to respiratory diseases akin to those caused by avian coronavirus and infectious bronchitis virus. The SARS-CoV encodes 23 putative proteins including four typical structural proteins: the spike, nucleocapsid, membrane, and envelope proteins.[4,7] These four proteins contribute to the host's immune response, as has been observed with many other transmissible viruses, such as gastroenteritis coronavirus, bronchitis virus, porcine respiratory coronavirus, and mouse hepatitis virus. In addition to the structural proteins, replicase 1a ORF, replicase 1b ORF and eight novel ORFs have been identified.

Eight SARS-CoV–encoded proteins have been shown to induce apoptosis.[9] These suggest that apoptosis may play an important role in helping with virus dissemination *in vivo*, minimising the inflammatory reaction and evasion of the host's defence mechanisms. Studies of its pathology have revealed diffuse alveolar damage as the most notable feature in persons who died of SARS. In SARS patients, apoptosis occurs in the alveolar epithelial cells. Four SARS-CoV accessory proteins have been shown to induce apoptosis, ORF-3a (also known as U274, SARS X1, or ORF-3), ORF-3b (also known as ORF-4), ORF-7a (also known as U122, SARS X4, or ORF-8) and ORF-8a.[10-12] We report the characterisation of another SARS-CoV group-specific gene product encoded by ORF-6 (also known as X3, ORF-7), which contains 63 amino acids (equivalent to nucleotides 27,074 - 27,265 in Tor2 genome sequence).[4] This small protein has no significant sequence homology to other proteins, and has been identified as an ER/Golgi membrane localised protein.[13]

We report that (1) there was overexpression of ORF-6 induced apoptosis; (2) caspase-3 was activated in the presence of ORF-6; (3) apoptosis induced by ORF-6 and ORF-7a was blocked by caspase-3 inhibitor, z-DEVD, and the JNK (c-Jun N-terminal kinase) inhibitor; and (4) ORF-6 and ORF-7a up-regulated levels of ER chaperone protein and 94kDa glucose-regulated protein (GRP94). All these features suggest that in addition to ORF-3a, ORF-3b and ORF-7a, ORF-6 is a new player involved in SARS-CoV–induced apoptosis and both ORF-6 and ORF-7a share similar pathways to induce apoptosis.

Results and discussion

Overexpression of ORF-6 protein induces apoptosis

In the search for SARS proteins that could induce apoptosis, ORF-6 and ORF-7a were amplified by polymerase chain reaction, and restriction sites BamHI and XhoI were introduced at the 5' and 3' end, respectively. The fragments of ORF-6 and ORF-7a were sub-cloned into mammalian expression vector GFP-N1. GFP-ORF-6 was transfected into Vero E6 and COS-7 cells. We observed that ORF-6

Hong Kong Med J 2010;16(Suppl 4):S22-6

Department of Biology, Hong Kong University of Science and Technology
ZD Ye, CK Wong, P Li, Y Xie
Protein Science Laboratory of the Ministry of Education, Department of Biological Sciences and Biotechnology, Tsinghua University
ZD Ye, P Li

RFCID project number: 01030832

Principal applicant and corresponding author:
Dr Yong Xie
Department of Biology, Hong Kong University of Science and Technology, Hong Kong SAR, China
Tel: (852) 2358 7340
Fax: (852) 2358 1559
Email: boyxie@ust.hk

was able to induce apoptosis when overexpressed in Vero E6 and COS-7 cells (Fig 1). Approximately one third of the cells died when 2 μg of ORF-6 DNA was transfected into cells. The number of apoptotic cells increased when more ORF-6 DNA was transfected into the cells (Fig 1). The death rates were comparable to rates caused by the overexpression of Bax, a well-known pro-apoptotic member of the Bcl-family, and ORF-7a, a SARS protein that has been shown to induce apoptosis.[11] GFP-N1 served as the negative control, as the transfection of 5 μg of GFP-N1 DNA did not induce apoptosis. These results showed that the ORF-6 protein did induce apoptosis.

Overexperession of ORF-6-induced apoptosis is caspase-3 dependent

ORF-7a induces apoptosis via a caspase-3–dependent pathway.[11] To determine if cell death induced by ORF-6 is also caspase-3 dependent, a caspase-3–specific inhibitor, z-DEVD, was used to block caspase-3 activation in the cells. In the absence of z-DEVD, approximately 60% of the cells underwent apoptosis when ORF-6 and ORF-7a were transiently transfected into the Vero E6 cells. However, when ORF-6 or ORF-7a were overexpressed in z-DEVD pretreated cells, the percentage of apoptotic cells was significantly decreased to approximately 20% (Fig 2a). In

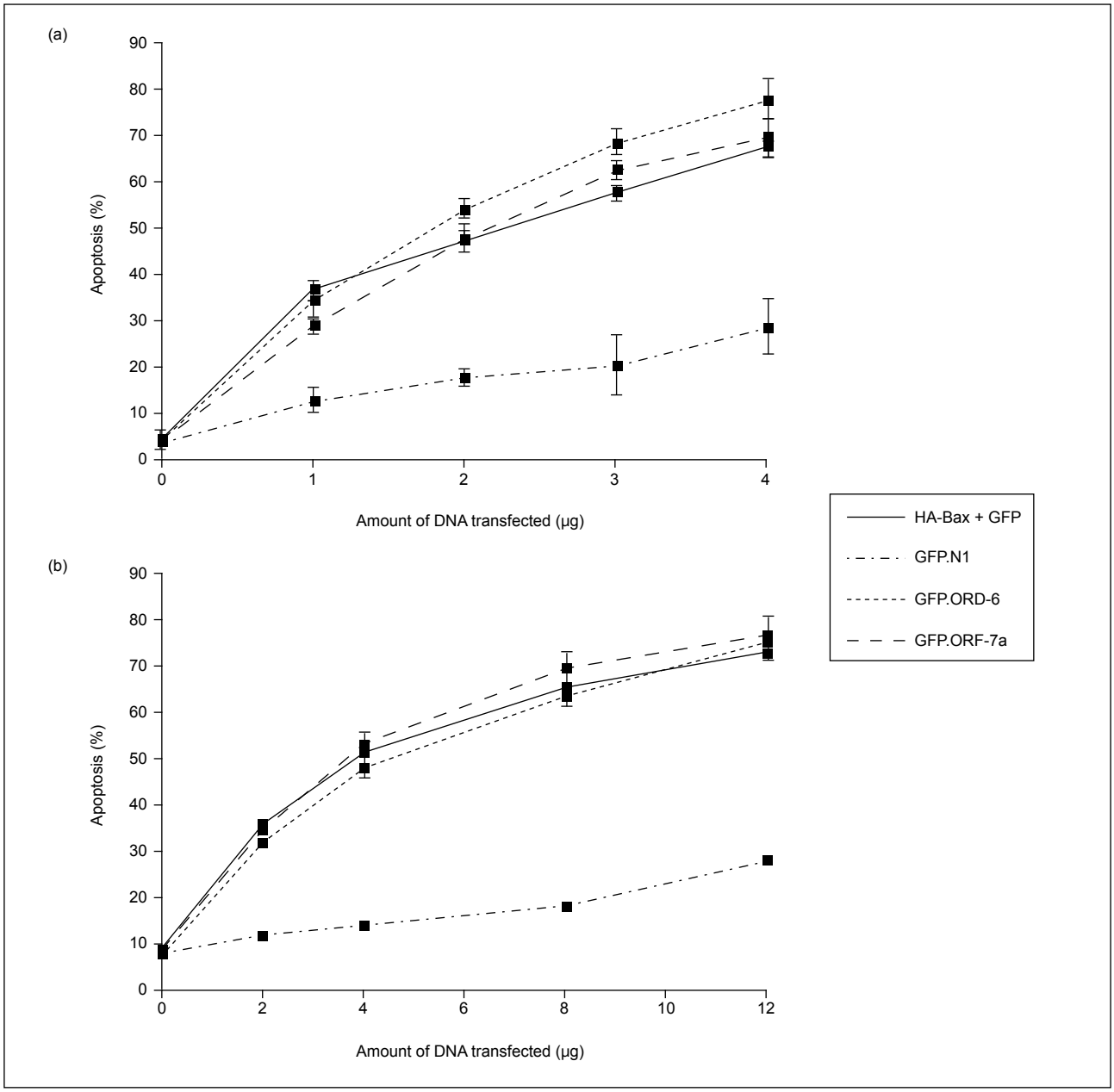


Fig 1. Overexpression of ORF-6–induced apoptosis in different cell lines
(a) COS-7 and (b) Vero E6 cells were transiently transfected with GFP and HA-Bax, GFP-N1, GFP-ORF-6 and GFP-ORF-7a for 24 hours. The nuclei of the cells were stained by Hoechst for 15 minutes. The number of healthy cells was counted under fluorescence microscopy with no DNA condensation and fragmentation. The percentage of apoptotic cells was calculated by the number of healthy cells over the total number of transfected cells. Experiments were repeated three times. Standard deviations are shown

parallel, caspase-3 activities were also monitored when the Vero E6 cells were transfected with ORF-6. The activity of caspase-3 can be detected by the amount of the 17kD active form of caspase-3 in cells by using caspase-3–specific antibodies. Both ORF-6 and ORF-7a induced caspase-3 activation, as the 17kD active form of caspase-3 was detected in both (Fig 2b). Cells without any treatment served as a negative control, and those treated with taurosporine served as a positive control. Similar results were obtained using 293T or COS-7 cells (data not shown). Our observation suggests that overexpression of ORF-6 induced apoptosis via a caspase-3–dependent pathway.

JNK inhibitor blocks ORF-6 and ORF-7a–induced apoptosis

One of the mechanisms for SARS protein–induced cell death is via the JNK pathway; JNK is phosphorylated in SARS-CoV–infected Vero E6 cells, and the JNK inhibitor (SP600125) can block SARS-CoV–infected Vero E6 cell–induced apoptosis. We investigated whether the JNK inhibitor could block ORF-6– and ORF-7a–induced apoptosis in Vero E6 cells. Interestingly, apoptosis induced by ORF-6 and ORF-7a was blocked by the JNK inhibitor. The blocking efficiency was similar to that of z-DEVD, as only approximately 10% of cells underwent apoptosis (Fig 2c). The JNK inhibitor was also able to block ORF-6 and ORF-7a in 293T and COS-7 cells (data not shown). These results suggest that the JNK inhibitor was able to block overexpression of ORF-6- or ORF-7a–induced apoptosis.

ORF-6 and ORF-7a induce ER stress

ORF-6 is localised in the ER/Golgi membrane, which is consistent with our immunostaining result for ORF-6 and ORF-7a in COS-7 cells (data not shown).[13] Since both ORF-6 and ORF-7a are ER localised, we suspect that they induce apoptosis through the ER stress pathway.

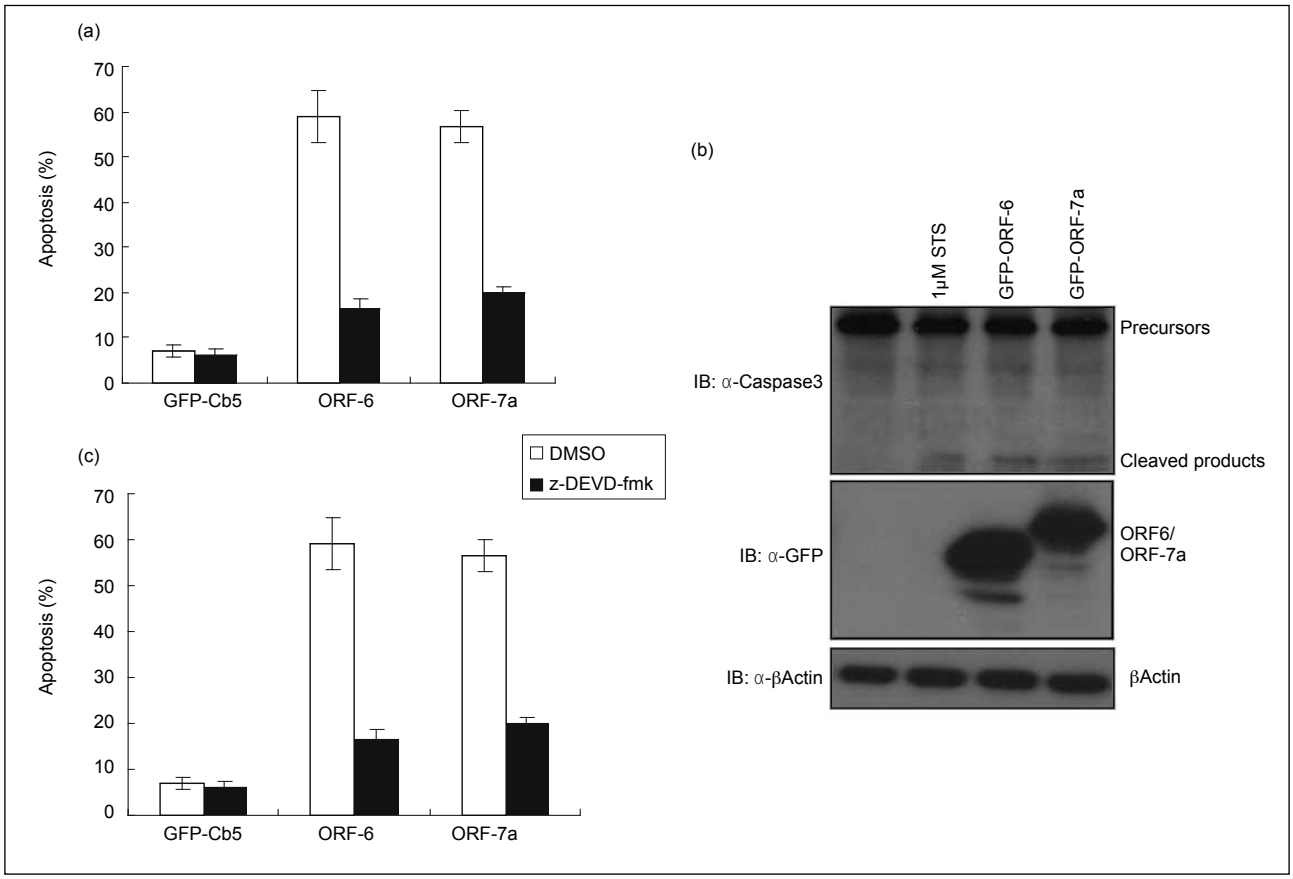


Fig 2. Overexpression of ORF-6–induced apoptosis is caspase-3 and JNK dependent

(a) Caspase-3 inhibitor (z-DEVD) blocks ORF-6–induced apoptosis. Vero E6 cells were incubated with either dimethyl sulfoxide or 50 μM z-DEVD-fmk for 30 minutes before they were transiently transfected with 3 μg of GFP-Cb5, GFP-ORF-6 and GFP-ORF-7a for 24 hours. The nuclei of the cells were stained by Hoechst for 15 minutes. The number of healthy cells was counted under fluorescence microscopy with no DNA condensation and fragmentation. The percentage of apoptotic cells was calculated by the number of healthy cells over the total number of transfected cells. Experiments were repeated three times and the standard deviations are shown. (b) Overexpression of ORF-6- and ORF-7a induces caspase-3 activation. Vero E6 cells were transiently transfected with GFP-ORF-6 and GFP-ORF-7a for 24 hours. Cell lysates were normalised to 2 μg/ml by lysis buffer and subjected to Western Blot with α-GFP, α-caspase-3 and α-Actin. For positive controls, cells were treated with 1 μM of staurosporine for 8 hours. Cells without any treatment were served as negative control. (c) JNK inhibitor blocked ORF-6–induced apoptosis. Vero E6 cells were incubated with either dimethyl sulfoxide or 40 μM JNK inhibitor for 30 minutes before they were transiently transfected with 3 μg of GFP-Cb5, GFP-ORF-6 and GFP-ORF-7a for 24 hours. Cell counts were done as mentioned in (b)

GRP94 is an ER-resident molecular chaperone protein, for which the expression level is increased by ER-stress.[14] We compared the protein level of GRP94 in order to check whether ER stress occurred in cells transfected with ORF-6 and ORF-7a. In such cells, the GRP94 protein level increased compared to cells without transfection (Fig 3a). However, the protein level was lower than in cells treated with 1 μM Thapsigargin, a chemical specific for induction of ER stress. In parallel, we conducted dose-dependent experiments by transfecting different amounts of ORF-6 and ORF-7a into cells. Consistently, the endogenous GRP94 protein level increased when the amount of ORF-6 or ORF-7 increased (Figs 3b and 3c). Our observation suggests that overexpression of ORF-6 and ORF-7a could induce apoptosis via the ER-stress pathway.

SARS-CoV encodes 23 putative proteins and in this context eight novel ORFs have been identified. Five of these (ORF-3a, ORF-3b, ORF-6, ORF-7a and ORF-8a) induce apoptosis when they are overexpressed in cells.[10-12,15,16] However, the responsible signalling pathway remains elusive. It has been shown that in SARS-CoV–infected Vero E6 cells, JNK is phosphorylated and that apoptosis was inhibited by both JNK and PI3K inhibitors.[17] We found that the JNK inhibitor inhibits both ORF-6– and ORF-7a–induced apoptosis. We therefore determined that ORF-6 and ORF-7a in SARS-CoV may be responsible for inducing phosphorylation of JNK, which in turn leads to apoptosis.

Acknowledgments

The study was supported by the Research Fund for the Control of Infectious Diseases, Food and Health Bureau, Hong Kong SAR Government (#01030832). We thank Dr Zhenguo Wu from Hong Kong University of Science and Technology for GFP-ORF6 and GFP-ORF7a constructs.

References

1. Drosten C, Gunther S, Preiser W, et al. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. N Engl J Med 2003;348:1967-76.
2. Ksiazek TG, Erdman D, Goldsmith CS, et al. A novel coronavirus

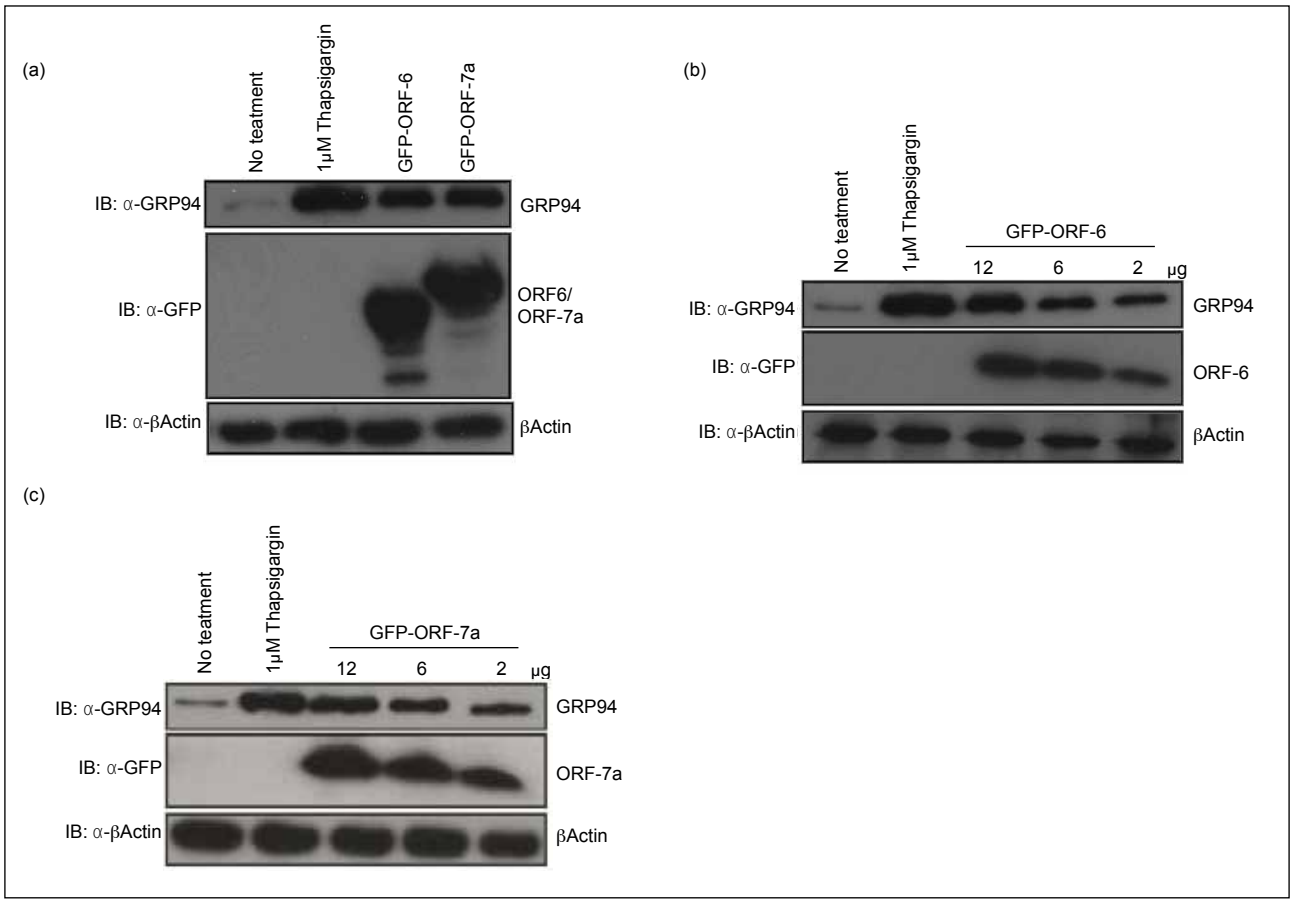


Fig 3. ORF-6 and ORF-7a induces ER stress

(a) Overexpression of ORF-6 and ORF-7a increases endogenous GRP94 protein level. Vero E6 cells were transiently transfected with 6 μg of GFP-ORF-6 and GFP-ORF-7a for 24 hours. Cell lysates were normalised to 1 μg/μl by lysis buffer and subjected to Western Blot with a-GFP, a-GRP94 and a-Actin. For positive controls, cells were treated with 1 μM of Thapsigargin for 24 hours. Cells without treatment served as negative controls. (b) and (c) ORF-6 and ORF-7a increase endogenous GRP94 protein level in dose-dependent manner. Vero E6 cells were transiently transfected with 2 μg, 6 μg and 12 μg of GFP-ORF-6 or GFP-ORF-7a for 24 hours. Cell lysates were normalised to 1 μg/μl by lysis buffer and subjected to Western Blot with α-GFP, α-GRP94 and α-Actin. For positive controls, cells were treated with 1 μM of Thapsigargin for 24 hours. Cells without treatment served as negative controls

associated with severe acute respiratory syndrome. N Engl J Med 2003;348:1953-66.
3. Lee N, Hui D, Wu A, et al. A major outbreak of severe acute respiratory syndrome in Hong Kong. N Engl J Med 2003;348:1986-94.
4. Marra MA, Jones SJ, Astell CR, et al. The Genome sequence of the SARS-associated coronavirus. Science 2003;300:1399-404.
5. Peiris JS, Chu CM, Cheng VC, et al. Clinical progression and viral load in a community outbreak of coronavirus-associated SARS pneumonia: a prospective study. Lancet 2003;361:1767-72.
6. Poutanen SM, Low DE, Henry B, et al. Identification of severe acute respiratory syndrome in Canada. N Engl J Med 2003;348:1995-2005.
7. Rota PA, Oberste MS, Monroe SS, et al. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science 2003;300:1394-9.
8. Tsang KW, Ho PL, Ooi GC, et al. A cluster of cases of severe acute respiratory syndrome in Hong Kong. N Engl J Med 2003;348:1977-85.
9. Tan YJ, Lim SG, Hong W. Regulation of cell death during infection by the severe acute respiratory syndrome coronavirus and other coronaviruses. Cell Microbiol 2007;9:2552-61.
10. Fielding BC, Tan YJ, Shuo S, et al. Characterization of a unique group-specific protein (U122) of the severe acute respiratory syndrome coronavirus. J Virol 2004;78:7311-8.
11. Tan YJ, Fielding BC, Goh PY, et al. Overexpression of 7a, a protein specifically encoded by the severe acute respiratory syndrome coronavirus, induces apoptosis via a caspase-dependent pathway. J Virol 2004;78:14043-7.
12. Chen CY, Ping YH, Lee HC, et al. Open reading frame 8a of the human severe acute respiratory syndrome coronavirus not only promotes viral replication but also induces apoptosis. J Infect Dis 2007;196:405-15.
13. Frieman M, Yount B, Heise M, Kopecky-Bromberg SA, Palese P, Baric RS. Severe acute respiratory syndrome coronavirus ORF6 antagonizes STAT1 function by sequestering nuclear import factors on the rough endoplasmic reticulum/Golgi membrane. J Virol 2007;81:9812-24.
14. Zhang K, Kaufman RJ. Signaling the unfolded protein response from the endoplasmic reticulum. J Biol Chem 2004;279:25935-8.
15. Law PT, Wong CH, Au TC, et al. The 3a protein of severe acute respiratory syndrome-associated coronavirus induces apoptosis in Vero E6 cells. J Gen Virol 2005;86:1921-30.
16. Yuan X, Shan Y, Zhao Z, Chen J, Cong Y. G0/G1 arrest and apoptosis induced by SARS-CoV 3b protein in transfected cells. Virol J 2005;2:66.
17. Mizutani T, Fukushi S, Murakami M, et al. Tyrosine dephosphorylation of STAT3 in SARS coronavirus-infected Vero E6 cells. FEBS Lett 2004;577:187-92.

Y Guan 管 軼

Genetic characterisation of H5N1 viruses isolated from different regions of outhern China

Key Messages

1. H5N1 viruses in different types of poultry in southern China and Southeast Asia have spread throughout South East Asia, Europe and Africa.
2. Multiple sublineages of H5N1 viruses were established in different regions. These sublineages were genetically and antigenically distinct, for which reason and based on our study, the World Health Organization selected different pandemic vaccine candidates.
3. Based on the recognition of multiple regional sublineages, a prediction of H5N1 spread to Europe was proposed.
4. Phylogenetic analysis revealed that Gs/Gd-like H5N1 was most likely derived from a low pathogenic H5 virus strain from migratory birds.
5. Some H5N1 novel reassortants contain internal genes that are only closely related to viruses isolated during 1970s, which suggests an unnatural evolutionary pathway.

Introduction

Preparing for a possible pandemic caused by highly pathogenic avian influenza A subtype H5N1 has been a global priority. The spread of the virus throughout Eurasia and Africa highlights the need for continued surveillance of influenza A virus in wild bird populations and poultry markets.

From July 2004 to June 2006, surveillance of poultry was carried out in live poultry markets in six provinces of southern China. From July 2005 to June 2006, of 53 220 poultry specimens, 2.4% were H5N1 positive (chickens, 0.5%; ducks, 3.3%; geese, 3.5%). This prevalence was higher than that determined from July 2004 to June 2005 (overall, 0.9%; chickens, 0.2%; ducks, 1.3%; geese, 2.0%). A winter-seasonal peak was observed throughout the surveillance period.[1]

Phylogenetic analysis of H5N1 viruses isolated from migratory birds showed that this virus is vectored by migratory birds. Identification of the first outbreak of H5N1 virus in migratory waterfowl from Qinghai lake and identification of regionally distinct sublineages showed that this virus had spread from its established source in southern China to other regions through poultry transport and bird migration (data not shown).[2,3]

Through the characterisation of low-pathogenic avian influenza H5 subtype viruses isolated from poultry and migratory birds in southern China and Europe from 1970s to 2000s, the dynamic nature of the influenza gene pool was revealed. Phylogenetic analysis revealed that Gs/Gd-like virus was probably derived from a low-pathogenic H5 virus in migratory birds. Phylogenetic analysis revealed frequent transmission between the eastern and western ends of the Eurasian continent that contributed to the spread of highly pathogenic H5N1 influenza A viruses (data not shown).[4]

Discussion

Long-term influenza surveillance in southern China enabled understanding of the genesis and transmission of this highly pathogenic H5N1 virus, and provided information for the World Heath Organization to improve global pandemic preparedness plan. This may help avert avian-to-human transmission, and thereby reduce the chance of viral establishment in humans and human-to-human transmission. The importance of animal and human influenza surveillance operations in affected regions deserves highlighting.

Acknowledgment

This study was supported by the Research Fund for the Control of Infectious Diseases, Food and Health Bureau, Hong Kong SAR Government (#01031022).

References

1. Smith GJ, Fan XH, Wang J, et al. Emergence and predominance of an H5N1 influenza variant in China. Proc Natl Acad Sci U S A 2006;103:16936-41.
2. Chen H, Smith GJ, Zhang SY, et al. Avian flu: H5N1 virus outbreak in migratory waterfowl.

Hong Kong Med J 2010;16(Suppl 4):S27-8

State Key Laboratory of Emerging Infectious Diseases, Department of Microbiology, The University of Hong Kong
Y Guan

RFCID project number: 01031022

Principal applicant and corresponding author:
Prof Yi Guan
State Key Laboratory of Emerging Infectious Diseases, Department of Microbiology, The University of Hong Kong, Faculty of Medicine Building, 21 Sassoon Road, Pokfulam, Hong Kong SAR, China
Tel: (852) 2819 9829
Fax: (852) 2817 1958
Email: yguan@hkucc.hku.hk

Nature 2005;436:191-2.
3. Chen H, Smith GJ, Li KS, et al. Establishment of multiple sublineages of H5N1 influenza virus in Asia: implications for pandemic control. Proc Natl Acad Sci U S A 2006;103:2845-50.
4. Duan L, Campitelli L, Fan XH, et al. Characterization of low-pathogenic H5 subtype influenza viruses from Eurasia: implications for the origin of highly pathogenic H5N1 viruses. J Virol 2007;81:7529-39.

MHL Ng 吳香玲
SH Cheng 鄭淑恆
KM Lau 劉建盟
GM Leung 梁卓偉
US Khoo 邱瑋璇
BCW Zee 徐仲鎂
JJY Sung 沈祖堯

Immunogenetics in SARS: a case-control tudy

Key Messages

1. Human leukocyte antigen (HLA) genotypes from 102 SARS patients (susceptible) and 108 SARS contacts (resistant) were obtained.
2. Allelic frequencies of the Class I (HLA-A, -B, and -Cw) and Class II (HLA-DR and -DQ) genes from these genetically unrelated subjects were compared.
3. A significantly higher frequency of DRB4*01010101 was found in the SARS-susceptible than SARS-resistant group. In contrast, significantly higher frequencies of HLA-B*1502 and HLADRB3*030101 were found in the SARS-resistant than SARS-susceptible group. However, none of these associations was significant after Bonferroni correction. Further, analysis of 10/36 genetically related families did not reveal any HLA alleles associated with SARS susceptibility or resistance.
4. We could not confirm previous findings of an HLA association with susceptibility or resistance to SARS. The significance of these associations needs to be validated by further independent studies.

Hong Kong Med J 2010;16(Suppl 4):S29-33

Prince of Wales Hospital, The Chinese University of Hong Kong:
Department of Anatomical and Cellular Pathology
MHL Ng, SH Cheng, KM Lau,
Department of Medicine and Therapeutics
JJY Sung
Centre for Clinical Trials
BCW Zee
Queen Mary Hospital, The University of Hong Kong:
Department of Medicine
GM Leung
Department of Pathology
US Khoo

RFCID project number: 02040412

Principal applicant and corresponding author:
Prof Margaret HL Ng
Department of Anatomical and Cellular Pathology, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, NT, Hong Kong SAR, China
Tel: (852) 2632 2179
Fax: (852) 2637 6274
Email: margaretng@cuhk.edu.hk

Introduction

Severe acute respiratory syndrome (SARS) was first reported in Guangdong province, China, in November 2002. Subsequently, multiple outbreaks occurred globally in early 2003, affecting over 8000 individuals of whom 700 died.[1,2] This highly infectious disease was caused by a novel coronavirus (SARS-CoV).[3] The clinical course was unpredictable and marked by a high fatality rate.[1] Its clinical severity varied from a febrile condition with mild respiratory symptoms and radiological evidence of lung involvement to severe respiratory distress with extensive lung damage and recourse to assisted ventilation.[1] In addition, individual susceptibility to the infection by the SARS-CoV was also variable. In SARS families, the index patient usually had one or more close household contacts (sharing rooms, meals and toilet facilities) with other family members and could have spread the virus in the early phase of the disease. Although family members were exposed to the same household environments, not all of them contracted SARS. This suggests a differential susceptibility to infection among family members. The mechanism of this variation remains unclear.

The T-cell system is a critical effector pathway for killing viruses by our body defences. To facilitate effective killing of the correct viral targets, human leukocyte antigens (HLA), as cell surface proteins, participate in selection and establishment of an antigen specific T-cell repertoire and in subsequent activation of such T cells during initiation of immune responses. Classes I (HLA-A, -B, and -Cw) and II (HLA-DR and -DQ) of the HLA antigens are responsible for CD8- and CD4-T cell responses, respectively. They are highly polymorphic and different alleles specify different functional significances in immune responses. The polymorphism in this machinery may contribute to the variations of patient susceptibility to immune-mediated/controlled diseases.

We examined HLA genotypes of 90 serologically confirmed SARS patients and found a significant increase in HLA-B*0703 frequency and a decrease in HLA-DRB1*0301 frequency in the SARS patients compared to the local Chinese population.[6] A similar study in Taiwan compared the HLA genotypes of 33 probable SARS patients with the normal population.[5] No significant association was identified except when comparing a highly selected group of five patients admitted to the intensive care unit with non-SARS health care workers; the HLA-B46 allele was found associated with this small SARS patient group. However, association with HLA-B46 was not found in our study.

Environmental factors such as contact history with SARS patients, prevention practices and personal hygiene are also important determinants for SARS infection. Thus, stringent defintion of close contacts is critical for further assignment to susceptible and resistant groups.

Methods

Study design

A case-control study was conducted from April 2005 to April 2007 in 536 SARS index patients and contacts recruited from the Hong Kong Department of Health SARS database (817 SARS patients and 936 contacts), which was established for contact tracing during the SARS outbreak. Verbal informed consent was obtained from each participant over the phone in collaboration with The University of

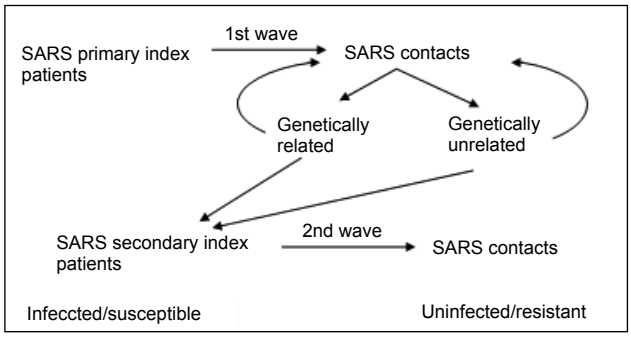


Fig 1. SARS families: human leukocyte antigen markers for differential susceptibilities with similar exposures

Hong Kong SARS research team. Of 433 saliva specimens collected, only 300 subjects including 176 from 36 genetically related (n=90) and 36 genetically unrelated (n=86) families completed relationship information or had adequate DNA extracted for HLA typing. The SARS index patients were considered as cases (susceptible) and contacts as controls (resistant) [Fig 1]. SARS was serologically confirmed and defined in accordance with the World Health Organization criteria.[1] A close contact was defined as a person who had cared for, lived with (in the same household), or came into direct contact with body fluids of the SARS patients within 10 days before hospital admission. The HLA allelic frequencies between these susceptible and resistant groups were compared for identification of positive and negative HLA risk markers by exclusion of all genetically related members. Genetically related family units within index contact pairs were also pulled out for family analysis with the aim to cross validate any positive observations.

Genetically unrelated case-control analysis

To prevent potential bias, all 90 cases and contacts from the 36 genetically related families were excluded. Thus, to identify HLA susceptible and resistant alleles, 102 SARS patients and 108 contacts were included for comparison of HLA allelic frequencies for the statistical significance of any association.

Genetically related familial analysis

Of 72 families recruited, only 36 were genetically related (Fig 2). The members (n=86) of the other 36 genetically unrelated families were husband and wife, co-inhabitants (boyfriends and girlfriends) or in-laws (not shown). The family relationship of the recruited subjects (n=90) among the 36 genetically related families is shown (Fig 2). Among these 36 related families, 10 were nuclear with both parents and ≥1 children (Fig 2). They were analysed statistically using the general family-based association test (FBAT), defined by $U=\Sigma$ Tij (Xij-E(Xij|Si)) (1), where i indexes pedigree, j indexes non-founders in the pedigree, and summation is over all i and j; Tij is a coding function for the trait of interest, and Xij is a coding function for the genotype. The coded genotype was chosen to reflect the selected mode of inheritance; for example, additive, dominant and recessive. Under the null hypothesis, the expected marker score, E(Xij|Si), was computed conditionally on the sufficient statistic 10, which is denoted by S. Under the null hypothesis, Mendel's laws determine which marker alleles are transmitted to the affected offspring. The FBAT compares the observed number of alleles that are transmitted with those expected in Mendelian transmissions. The assumption of Mendelian transmissions is all that is needed to ensure valid results of the FBAT approach. An excess of alleles of one type among the affected indicates that a disease-susceptibility locus for a trait of interest is linked and associated with the marker locus.

Human leukocyte antigen typing

DNA was extracted from the saliva specimens from all SARS patients and contacts. Sequencing based typing (SBT) of HLA-A, -B, -Cw, DRB and DQB1 was performed according to the protocols established by the International Histocompatibility Working Group (IHWG). Sequencing data were analysed using the SBTengine software (Genome Diagnostics, Netherlands). Dynal AllSet+ SSP kit was used for DQA1 typing (Invitrogen, UK). The SBT method based on direct sequencing of the specific HLA locus after PCR-amplification has been established as one of the most reliable and sensitive HLA genotyping techniques. Locus-specific PCR amplification was first performed with primer sets as suggested by IHWG, using Taq polymerase (Promega) on 9700 thermal cycler (Applied Biosystems). Specific PCR products were excised from gels after electrophoresis and purified by gel DNA extraction kit (Qiagen). Sequencing reactions were then performed using BigDye 3.1 reagent (Applied Biosystems) and resolved on a 3130 sequencer (Applied Biosystems). Finally, the sequencing data were analysed by sequence alignment and database matching using the SBTengine software (Genome Diagnostics).

Results

Comparison between case (susceptible) and control (resistant) groups

A total of 102 SARS patients and 108 contact non-SARS subjects were included. The SARS group consisted of 64 males and 38 females aged 13 to 88 (mean, 42) years. The contact non-SARS group consisted of 62 males and 46 females, with age information unavailable. We compared the HLA frequencies between the two groups. A significantly higher frequency of DRB4*01010101 was found in the SARS than contact group (50.98% vs 30.56%, p=0.0031, Pc=0.093, odds ratio [OR]=2.364, 95% confidence interval [CI]=1.344 to 4.156), whereas significantly higher frequencies of HLA-B*1502 (25.93% vs 9.80%, p=0.0037, Pc=0.2035, OR=0.3106, 95% CI=0.1421 to 0.6788) and HLA-DRB3*030101 (27.78% vs 14.71%, p=0.0282, Pc=0.846, OR=0.4483, 95% CI=0.2246 to 0.8948) were observed in the contact than SARS group (Table 1). However, none of these associations was significant after the Bonferroni correction. In addition, in contrast to our

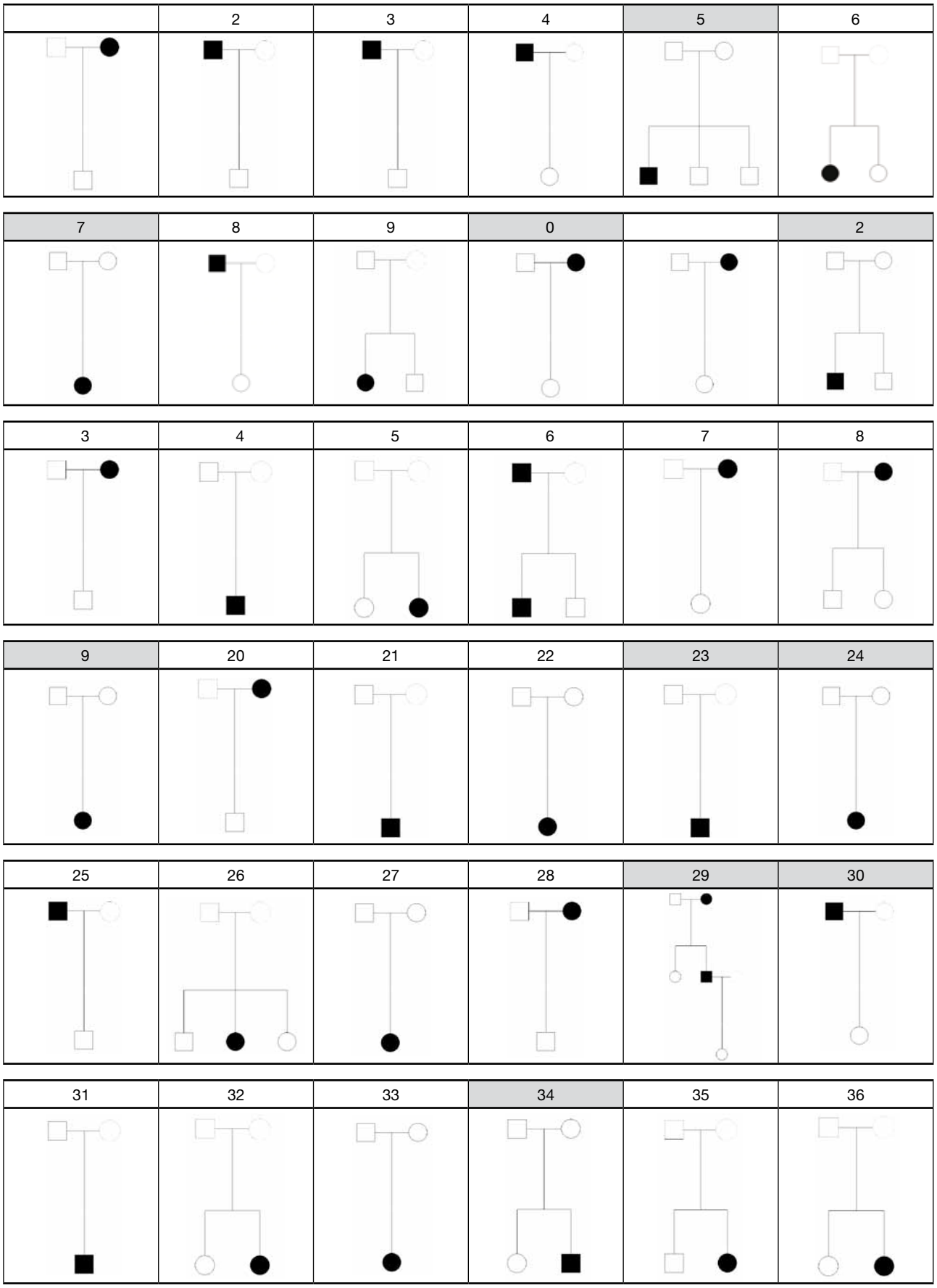


Fig 2. Summary of the pedigrees of the 36 genetically related families
Black boxes/circles indicate SARS index patients, whereas open boxes/circles denote resistant contacts

Table 1. Associations of human leukocyte antigen genes with SAR

	No. (%) of SARS (n=102)	No. (%) of contacts (n=108)	P value	P value (corrected)	Odds ratio	95% CI
B*1502	10 (10)	28 (26)	0.004	0.204	0.311	0.142 to 0.679
DRB3*030131	15 (15)	30 (28)	0.028	0.448	0.448	0.225 to 0.895
DRB4*01010101	52 (51)	33 (31)	0.003	0.093	2.364	1.344 to 4.156

Table 2. Data analysis of the 12 alleles with threshold frequencies by family-based association test (FBAT)

Allele	Freq	HW	Freq parent	HW parents	Nbr info fam	P value (FBAT)	P value (FBATI)	Power (FBAT)	Power (FBATI)	P value (Wald)	P value (WaldI)	Affection Status
A*020301	0.385	0.079	0.472	0	9	0.564	1	0.051	0.05	1	1	No SARS
A*110101	0.615	0.079	0.528	0	9	0.564	1	0.051	0.05	1	1	No SARS
Cw*0403	0.500	0.002	0.500	0	9	1.000	1	0.050	0.05	1	1	No SARS
Cw*040101	0.500	0.002	0.500	0	9	1.000	1	0.050	0.05	1	1	No SARS
B*1521	0.500	0.002	0.500	0	9	1.000	1	0.050	0.05	1	1	No SARS
B*1527	0.500	0.002	0.500	0	9	1.000	1	0.050	0.05	1	1	No SARS
DRB1*0404	0.500	0.002	0.481	0	9	1.000	1	0.050	0.05	1	1	No SARS
DRB1*140101	0.500	0.002	0.519	0	9	1.000	1	0.050	0.05	1	1	No SARS
DRB4*01010101	0.500	0.002	0.557	0	9	1.000	1	0.050	0.05	1	1	No SARS
DRB3*0201	0.500	0.002	0.443	0	9	1.000	1	0.050	0.05	1	1	No SARS
DQB1*030201	0.500	0.002	0.500	0	9	1.000	1	0.050	0.05	1	1	No SARS
DQB1*0502	0.500	0.002	0.500	0	9	1.000	1	0.050	0.05	1	1	No SARS

previous observations, no HLA-B*0703 was identified in either the SARS or contact group, and HLA-DRB1*0301 was observed in equal frequency in both groups.

Analysis of genetically related family units

Ten nuclear families from the 36 genetically related families were analysed using the general FBAT. From this calculation, 12 HLA alleles with significant threshold frequency were further analysed but none was found to be significantly associated (Table 2).

Discussion

Clinical outcome of viral infections is an interactive result of the virus and the host immunogenetic response; HLA is one important immunogenetic determinant that may attribute to the diversity of responses and outcomes.

Our previous exploratory HLA study had revealed a positive association of HLA-B*0703 and negative association of HLA-DRB1*0301 with SARS group. Coinheritance of HLA-B*0703 with HLA-B60 was found to confer an even higher risk for SARS.[4] There were several limitations of the aforementioned study. First, comparison of the HLA genotypic frequencies was made between the SARS group and the normal population, which had not been exposed to SARS virus. Second, the normal population data for comparison were published in 1997 based mainly on a low-resolution platform including serotyping. Third, the sample size was relatively small. Our current study entailed a number of improvements. First, the control group could actually be viewed as SARS resistant, unlike the normal population, as they were exposed to the SARS virus and the circumstances of exposure were similar to those of the index SARS patients. Thus, more subtle associations if present could be detected. Second, a higher-resolution platform was employed and the associations if found would yield more refined information. Third, HLA-Cw was also typed for the first time, for potential association with this set of Class I genes.

Our study could not confirm the previous findings of HLA associations using an additional independent SARS cohort with improved design and typing methods. Thus, at this stage the previous findings remain not validated. The reason for the discrepancy may hinge on the sampling and the typing method.

Although positive and negative associations were found with one and two specific HLA genotypes respectively, their significance needs to be validated by further independent studies. Our current findings do not suggest a strong involvement of HLA with the genetic susceptibility to SARS, at least when based on the HLA genes that we studied. Negative findings were also reported in a study performed in China. Other HLA genes or alternative immunogenetic response genes may play a role in this aspect and deserve exploration.

Acknowledgement

This study was supported by the Research Fund for the Control of Infectious Diseases, Food and Health Bureau, Hong Kong SAR Government (#02040412).

References

1. World Health Organization. Severe acute respiratory syndrome (SARS), 2003. Available at: http://www.who.int/csr/sars/en/index.html.
2. Lee N, Hui D,Wu A, et al. A major outbreak of severe acute respiratory syndrome in Hong Kong. N Engl J Med 2003;348:1986–94.
3. Peiris JS, Lai ST, Poon LL, et al. Coronavirus as a possible cause of

severe acute respiratory syndrome. Lancet 2003;361:1319–25.
4. Ng MH, Lau KM, Li L, et al. Association of human-leukocyte-antigen class I (B*0703) and class II (DRB1*0301) genotypes with susceptibility and resistance to the development of severe acute respiratory syndrome. J Infect Dis 2004;190:515-8.
5. Lin M, Tseng HK, Trejaut JA, et al. Association of HLA class I with severe acute respiratory syndrome coronavirus infection. BMC Med Genet 2003;4:9.

WH Seto 司徒永康

Simultaneous detection of SARS coronavirus and influenza A viruses using real-time polymerase chain reaction

Key Messages

1. Simultaneous detection of SARS-CoV and influenza A virus is described.
2. The rapid diagnostic method is less expensive than other methods.
3. Automation of molecular diagnosis is possible.

Introduction

In early 2003, severe acute respiratory syndrome (SARS) was recognised as a newly emerging pneumonic disease and a novel coronavirus, named SARS coronavirus (SARS-CoV), was implicated as the causative agent of the epidemic. In Hong Kong, the outbreak affected more than 1700 people with 290 deaths up to July 2003. Specific laboratory tests to detect viral RNA and antibody responses were used to establish a cause in patients suspected to have SARS. As one of the World Health Organization SARS network laboratories, the Department of Microbiology, Queen Mary Hospital, The University of Hong Kong provided rapid laboratory diagnoses (within 24 hours) for clinically suspected SARS cases using a polymerase chain reaction (PCR) technology. The first-generation PCR protocol was found to be highly specific but insensitive for rapid diagnosis of SARS-CoV infection.

During the latter part of 2003, a viral influenza A epidemic in North America and Europe caused considerable morbidity and mortality. The overlapping of SARS and the influenza season in Hong Kong created additional difficulty for establishing an early clinical diagnosis. The situation became more complicated when fatal cases of avian influenza H5N1 were reported in Southeast Asia. Rapid laboratory confirmation of both SARS-CoV and influenza virus is important not only to facilitate timely therapeutic interventions and clinical management, but also to identify the beginning of any new influenza or SARS epidemic.

In this project, a highly sensitive multiplex real-time PCR was developed for simultaneous rapid diagnosis of both SARS-CoV and influenza A viruses (including human seasonal influenza [eg H3N2] and avian influenza H5N1 causing epidemics in Southeast Asia). It included a retrospective analysis of 500 stored respiratory samples collected from confirmed cases of SARS or viral influenza infection in 2003, and another 500 samples recruited for prospective study in 2006.

Methods

A total of 1150 nasopharyngeal aspirates were collected from patients suffering respiratory tract infections or flu syndrome between 2003 and 2007. Conventional laboratory diagnosis for influenza A and B viruses, parainfluenza virus types 1, 2, and 3, respiratory syncytial virus, and adenovirus were performed using an immunofluorescence, commercial screening kit and viral culture on MDCK, HepII and LLCMK2 cells. Patient serum samples were also collected for serologic testing for SARS-CoV. RNA extraction was performed automatically by QIAamp Virus BioRobot 9604 (Qiagen, Hilden, Germany) for total viral nucleic acid load (Table 1). Initial processing of specimens was performed under biohazard level-2 containment. According to the manufacturer's instructions, a sample volume of 220 μL was mixed with 240 μL of AL buffer and 40 μL protease solution. The mixture was incubated at 60ºC for 10 minutes before transferring to the rack of the Robotic 9604 System containing 96 samples tubes. Nucleic acid precipitation was initiated by addition of 275 μL absolute ethanol

Hong Kong Med J 2010;16(Suppl 4):S34-6

Department of Microbiology, Queen Mary Hospital
WH Seto

RFCID project number: 03040112

Principal applicant and corresponding author:
Dr WH Seto
Department of Microbiology, Queen Mary Hospital, Pokfulam Road, Hong Kong SAR, China
Tel: (852) 2855 3206
Fax: (852) 2872 4555
Email: whseto@ha.org.hk

Table 1. Workflow and manipulation time for multiplex real-time polymerase chain reaction (PCR) assays

Workflow	Hands-on time for 32 samples
Decontamination and lysis	45 min
Automatic RNA extraction, purification and concentration (waiting time 40 min)	0 (The second batch of samples can be processed)
Setting master mix for multiplex real-time PCR	30 min
Real-time PCR (waiting time 45 min)	0 (The third batch of samples can be processed)
Turnaround time	2 hr 40 min

and 250 μL AW1 buffer followed by transferring to the QIAamp 96 well vacuum manifold. RNA extracted was washed once with 360 μL AW1 buffer and twice with 1 μL AW2 buffer. A final volume of 50 to 60 μL total nucleic acid was eluted by addition of 86 μL AVE elution buffer. The sequential steps of nucleic acid precipitation, washing and elution were operated automatically. Each run of nucleic acid extraction for 96 samples required 3 hours. A multiplex real-time PCR was designed against SARS-CoV (P gene) and influenza A (M gene) with confirmation by specific fluorescent labelled hybridisation probes.

Real-time PCR was performed using LightCycler PCR (Roche Diagnostics GmbH, Manheim, Germany) with the following conditions: 61ºC for 20 minutes, 95ºC for 30 seconds; amplification at 95ºC for 5 seconds, 55ºC for 15 seconds, and 72ºC for 13 seconds for 50 cycles; cooling at 40ºC for 30 seconds. Positive and negative controls were included in each run, the National Committee for Clinical Laboratory Standards for the molecular diagnosis of infectious diseases were adopted to prevent cross contamination.

Results

Among the 1150 samples tested, 323 positive samples were obtained from patients during the 2003 SARS epidemic (Table 2). No SARS-CoV was detected in samples obtained between 2004 and 2007. During the SARS epidemic in 2003, 638 paired (10-21 days apart) serum samples were available for parallel testing with 680 nasopharyngeal aspirate samples. Among 470 nasopharyngeal aspirate samples collected from late 2003 to 2007, only 65 paired samples for SARS-CoV serology assay were available due to early discharge of patients. For 405 single serum samples collected during the early onset of disease, a serology result of ≥1:20 dilution was interpreted as positive. Influenza A viruses (H1 and H3 subtypes) were detected among 275 samples and no H5 subtype was identified. All 278 samples positive for SARS-CoV by multiplex PCR for the P gene were confirmed by real-time PCR for the N gene on the same DNA extract. The resolved performance of real-time PCR was validated against conventional laboratory diagnosis with 86% and 93% sensitivity for SARS-CoV and influenza A viruses, respectively. The real-time PCR assay exhibited 100% specificity for the detection of both viral pathogens.

Discussion

Sero-diagnosis for SARS-CoV infection is reliable and specific, but sero-conversion can only be detected by about day 10 of the illness. When patients are treated with immunomodulator drugs such as steroids, the detection may be delayed until the third or fourth week of illness. The first-generation in-house RT-PCR assay has been shown to be highly specific but insensitive. In the previous study, manual RNA extraction was performed using 140 μL sample volume. In this study, the automatic system used a larger initial sample volume (220 μL) for viral RNA extraction, thus providing higher diagnostic sensitivity with a 30% saving of manpower. The multiplex real-time PCR assay also increased the specificity of this assay for the differential diagnosis of SARS-CoV and influenza A virus and was suitable for routine application.

Rapid diagnosis of SARS-CoV by PCR assays is useful during an epidemic. In the post-epidemic period, no SARS-

Table 2. Performance of multiplex real-time polymerase chain reaction (PCR) for rapid detection of SARS-CoV and influenza A viruses on nasopharyngeal aspirate specimens

No. of specimens tested	Sero- onversion of SARS-CoV	Conventional diagnosis of influenza A			No. of specimens positive for real-time PCR			
		H1	H3	H5	SARS-CoV	H1	H3	H5
323	+*	–	–	–	278†	0	0	0
186	–	–	+	–	0	0	177‡	0
89	–	+	–	–	0	79	0	0
897	–	–	–	–	0	0	0	0
264§	–	–	–	–	0	0	0	0

* 703 paired serum (10-21 days apart) and 405 single serum at early onset of disease were available. A 4-fold rise in paired serum titre was interpreted as positive. For single serum, a positive serology result for ≥1:20 dilution was interpreted as positive

† The 278 samples positive for the P gene were confirmed by N gene real-time PCR assay

‡ Two samples were positive for both influenza A and adenovirus

§ Samples positive for other viral pathogens included 12 parainfluenza type 1, two parainfluenza type 2, 39 parainfluenza type 3, 121 respiratory syncytial virus (RSV) and 87 adenovirus, five with co-infections (three RSV & parainfluenza type 1, two RSV & parainfluenza type 3)

CoV was detected by the multiplex real-time PCR assay. To monitor reemerging of SARS-CoV infection, rapid diagnosis is important not only for timely therapeutic interventions but also to identify the beginning of a new outbreak. Our study highlights the high throughput and performance of automatic RNA extraction and multiplex real-time PCR assay, which appear suitable for large-scale routine diagnosis in the event of future SARS and/or influenza epidemics.

Acknowledgements

This study was supported by the Research Fund for the Control of Infectious Diseases, Food and Health Bureau, Hong Kong SAR Government (#03040112). We thank the technical team of the Department of Microbiology, Queen Mary Hospital for sorting and proper storage of specimens.

References

1. Yam WC, Chan KH, Chow KH, et al. Clinical evaluation of real-time PCR assays for rapid diagnosis of SARS coronavirus during outbreak and post-epidemic periods. J Clin Virol 2005;33:19-24.
2. Poon LL, Chan KH, Wong OK, et al. Early diagnosis of SARS coronavirus infection by real time RT-PCR. J Clin Virol 2003;28:233-8.
3. Yam WC, Chan KH, Poon LL, et al. Evaluation of reverse transcription-PCR assays for rapid diagnosis of severe acute respiratory syndrome associated with a novel coronavirus. J Clin Microbiol 2003;41:4521-4.

KYK Chan 陳遠光
MS Xu 許美術
JCY Ching 程子忻
VS Chan 陳秀芳
YC Ip 葉瑩芝
L Yam 任燕珍
CM Chu 朱頌明
ST Lai 黎錫滔
KM So 蘇文傑
TY Wong 黃天祐
PH Chung 鍾沛康
P Tam 譚廣亨
SP Yip 葉社平
P Sham 沈伯松
CL Lin 林成龍
GM Leung 梁卓偉
JSM Peiris 裴偉士
US Khoo 邱瑋璇

Association of a single nucleotide polymorphism in the CD209 (DC-SIGN) promoter with SARS severity

Key Messages

1. In this genetic risk-association study involving about 1300 subjects, SARS patients carrying the DC-SIGN promoter -336G variant had lower risk of having higher lactate dehydrogenase levels on admission, an independent prognostic indicator for severity of SARS-CoV infection.
2. *In vitro* functional studies demonstrated that the DC-SIGN -336G promoter provided a less effective binding site and lower promoter activity, which may lead to reduced DC-SIGN protein expression and hence may contribute to a reduced immune-response with reduced lung injury during the progression of SARS infection.

Hong Kong Med J 2010;16(Suppl 4):S37-42

Li Ka Shing Faculty of Medicine, The University of Hong Kong:
Department of Pathology
KYK Chan, MS Xu, JCY Ching, YC Ip, US Khoo
Department of Surgery
VS Chan, P Tam, CL Lin
Department of Psychiatry
Sham P
School of Public Health
GM Leung
Department of Microbiology
JSM Peiris
Pamela Youde Nethersole Hospital
L Yam
United Christian Hospital
CM Chu
Princess Margaret Hospital
ST Lai, KM So
Infectious Disease Control, Hospital Authority
TY Wong
Department of Health, Hong Kong SAR
PH Chung
Department of Health Technology and Informatics, The Hong Kong Polytechnic University
SP Yip

RFCID project number: 04050252

Principal applicant and corresponding author:
Dr Ui-Soon Khoo
Room 324, University Pathology Building, Queen Mary Hospital, Pokfulam Road, Hong Kong SAR, China
Tel: (852) 2855 4410
Fax: (852) 2872 5197
Email: uskhoo@pathology.hku.hk

Introduction

Severe acute respiratory syndrome (SARS) is an acute respiratory disease resulting from infection by a novel coronavirus (SARS-CoV). Dendritic cell–specific ICAM-3 grabbing non-integrin (DC-SIGN) is a C-type lectin expressed on the surface of subsets of human DCs and alveolar macrophages, and functions both as a cell adhesion and as a pathogen recognition receptor to facilitate infection. Pseudo-particles presenting SARS-CoV spike protein can bind to DC-SIGN. Homozygous genotypes of the tandem-repeat polymorphism of DC-SIGNR, which share 77% similarity with DC-SIGN, are protective against SARS-CoV infection.

As a pathogen recognition receptor, DC-SIGN allows DCs to capture and interact with numerous pathogens, such as *Mycobacterium tuberculosis*, HIV-1 and dengue viruses, etc. As an adhesion receptor, it plays an important role in many DC functions, such as the DC–T-cell interaction and DC migration. Host genetic factors have been shown to result in differences in host susceptibility to SARS infection and outcome of the disease. These include the genetic polymorphisms in L-SIGN, MBL, HLA-B and HLA-DRB1 for host susceptibility; and polymorphisms in ACE1, ICAM3, MXA and FcγRIIA for disease outcome.

In the DC-SIGN promoter region, a putative functional single nucleotide polymorphism (SNP), -336A>G, (rs4804803) affecting an Sp1-like transcription factor binding site has been described.[1] This SNP was found to affect the transcription activity of DC-SIGN *in vitro*. It has significant association with susceptibility for HIV-1 and *M tuberculosis* infections, and with the severity of dengue.[1-3]

As the SARS-CoV spike proteins share great similarity with HIV envelope proteins, we hypothesised that the DC-SIGN promoter SNPs may also be associated with the severity of SARS-CoV infection. Therefore, we analysed the promoter SNPs of DC-SIGN for genetic association to SARS-CoV using a large case-control study, and performed *in vitro* functional studies to verify the effect of DC-SIGN promoter SNPs on promoter activity and its influence on DC-SIGN expression *in vivo*.

Methods

Study design

This case-control genetic association study was conducted from June 2006 to May 2008 to examine the contribution of DC-SIGN -336A/G SNP to SARS-CoV infection and/or association with SARS patients' clinico-pathological outcomes.

A total of 824 SARS patients confirmed by serology and/or RT-PCR for SARS were recruited from the Pamela Youde Nethersole Hospital, Princess Margaret Hospital, United Christian Hospital, Queen Mary Hospital, Alice Ho Miu Ling Nethersole Hospital, and Prince of Wales Hospital. The 471 controls included 281 household contacts (genetically unrelated household members of SARS patients) who remained unaffected and sero-negative, and 190 health care workers who had worked in SARS wards but remained disease-free and sero-

negative. All the subjects and controls were Chinese.

Main outcome measures

The clinical data of the SARS cases were retrospectively obtained from the Hospital Authority, with permission from all attending clinicians. The data included age, sex, length of hospital stay, extent of intensive care unit treatment, whether patients received assisted ventilation, any form of steroid treatment or intravenous immunoglobulin, as well as final outcomes in terms of survival and death. Results of haematological and biochemical laboratory investigations on admission included the haemoglobin level, absolute lymphocyte count, platelet count, white blood cell count; and biochemical indices of alanine-aminotranseferase, albumin, globulin, creatinine-kinase, lactate-dehydrogenase, urea, sodium, potassium and serum creatinine.

Study instruments

Genotyping for DC-SIGN -336A/G promoter SNP was performed using Allelic Discrimination TaqMan Assay (Applied Biosystems). Proper controls and replicates were included for quality control. DC-SIGN -336A/G genotyping results were also integrated with the ICAM3 +443T/C genotyping results[4] for combined genotype risk-association analysis.

In vitro functional studies of DC-SIGN -336A/G promoter SNP was performed using (1) an electrophoretic mobility shift assay (EMSA) and (2) a luciferase-reporter promoter activity assay. The EMSA was used to investigate the differential binding of nuclear extract to the -336A/G polymorphic site. Complementary oligonucleotide pairs harbouring either -336A or -336G of the DC-SIGN promoter were ^{32}P-end-labelled and were incubated with HeLa cell nuclear extracts. Consensus Sp1 and AP2 probes were used to compete with -336A or -336G probe for nuclear extract binding. Before incubation, unlabelled -336A or -336G oligonucleotide probes were added to compete for the radiolabelled probe reaction mixture at 25 or 50 times molar excess. The protein-DNA complexes were resolved using non-denaturing polyacrylamide gel electrophoresis.

The luciferase-reporter promoter activity assay was used to examine whether -336A or G allele would give rise to lower promoter activity. One μg of luciferase-reporter DC-SIGN promoter construct plasmid and 0.1 μg of Renilla plasmid were transfected into the HeLa cell lines using lipofectamine 2000 (Invitrogen). After transient transfection for 24 hours, the cells were harvested and the activity of the promoter constructs assayed using the dual-luciferase reporter assay system (Promega). To determine promoter activity, firefly luciferase-expression levels were normalised against Renilla-luciferase levels. The luciferase expression levels between pGL3-basic/DC-SIGN-336A and pGL3-basic/DC-SGIN-336G were compared. The pGL3-basic plasmid was used as a negative control. Light emission was measured; the experiments were performed as three sets of triplicates. Results were compared using the Student's *t* test.

The expression level of DC-SIGN in the peripheral blood mononuclear cells (PBMCs) derived DCs (which carried different -336A/G genotypes) was compared. $CD14^+$ monocytes were isolated from PBMCs and then cultured in the presence of GM-CSF and IL-4, and were harvested after 5 days of culture for DC-SIGN expression analysis by flow cytometry. The percentage and mean fluorescence level of positive cells were measured by counting the cells demonstrating a higher signal than the control.

Statistical analysis

For risk-association analysis, genotype distributions of the patients and controls were compared using the χ^2 test; strength of association was measured using an odds ratio (OR) and 95% confidence intervals (CI). Genotyping results were checked for Hardy-Weinberg equilibrium. For *in vitro* studies, statistical significance was calculated using the Student's *t* test. The χ^2 test was used to test for possible association with nominal clinical outcome measures. For numerical variables, each was first analysed by Student's *t* test. If significant, it was further stratified and then examined using the χ^2 test. Logistic regression analysis was applied to adjust for age and gender.

Results

Genetic susceptibility analysis

A total of 824 SARS cases and 471 controls were successfully genotyped. The SNP was in Hardy-Weinberg equilibrium in both patients and controls. There was no significant difference between patients and controls in terms of genotypes (P=0.93) or allele frequency (P=0.81). Trend of association was not observed. Therefore, further investigations with more healthy controls were not performed.

Severity association analysis

Because only three subjects were homozygous for GG polymorphism, we combined GG and GA heterozygous genotypes for severity association analysis. The mean±standard error of the mean of lactate dehydrogenase (LDH) levels in the AA group was significantly higher than that in GG/GA group (1.14±0.03 vs 1.03±0.04, P=0.019), suggesting that LDH levels were associated with the DC-SIGN-336A/G SNP. All other parameters tested (ie ALT, albumin, creatinine kinase, haemoglobin, lymphocyte, platelet, white blood cell and globulin) showed no significant association with the genotype groups by independent *t* test (Table 1). To examine the risk level, patients were categorised into two subgroups, namely high and low LDH levels (Table 2). The frequency of G-carriers in persons with low LDH levels was significantly higher than in those with high LDH levels (18.46% vs 8.33%, P=0.015, OR=0.40, 95% CI=0.19-0.85); the DC-SIGN-336G allele was overrepresented in the low LDH level group than in the high LDH level group (9.49% vs 4.17%,

P=0.014, OR=0.41, 95% CI=0.20-0.86). These results suggested that G carriers had a 2.44-fold greater chance of having lower LDH levels. The association remained significant after adjustment for age and gender as well as genetic relationships of the subjects.

DC-SIGN and ICAM3 combined genotype analysis

The genotyping results of DC-SIGN -336A/G SNP of the SARS patients was combined with our previously reported genotyping results of ICAM3 Asp143Gly (+443T/C) SNP[4] (Table 3). The wild-type genotypes of both SNPs (ie ICAM3+443TT/DC-SIGN-336AA genotypes) were of highest frequency in the cohort. Analysis of the mean LDH level on admission of the combined genotypes showed that combined genotype +443CC/-336AA was the highest. Non-parametric *t*-test of +443CC/-336AA genotype and the wild-type genotype groups showed significant difference in the LDH level (P=0.033). The frequency of patients with high LDH levels in each combined genotype was compared to that of the wild-type +443TT/-336AA genotype using the χ^2 test (Table 3). Significant association of the overall genotypes was observed (P=0.024). Comparison of the wild-type genotype with each combined genotype also showed association of the combined +443CC/-336AA genotype with higher LDH levels (P=0.021, OR=4.34, 95% CI=1.34-14.12), although the number of cases having the combined genotype was small (n=12). Multivariate logistic regression analysis for synergistic effect between the SNPs indicated that interaction between the two SNPs was not significant (P=0.845).

Electrophoretic mobility shift assay and promoter activity assays

DC-SIGN -336A/G was found to harbour the Sp-1 transcription factor binding site, with the - 336G allele harbouring the binding site, but being abolished with the

Table 1. Univariate association between laboratory parameters on admission and genotypes of C-SIGN-336A/G among SARS patients

Parameters	G-carrier (GG/GA)			Non G-carrier (AA)			P value*
	No.	Mean	SEM	No.	Mean	SEM	
Alanine aminotransferase	127	0.66	0.05	656	0.82	0.04	0.093
Albumin	127	40.19	0.38	655	39.78	0.17	0.337
Creatine kinase	133	133.22	16.76	659	162.03	14.47	0.384
Haemoglobin	134	13.53	0.12	687	13.28	0.07	0.090
Lymphocyte	134	0.94	0.03	683	0.92	0.02	0.501
Platelet	134	169.79	4.64	687	180.16	2.51	0.051
White blood cell	134	5.63	0.18	687	5.67	0.09	0.869
Globulin	93	34.76	0.48	537	33.87	0.21	0.107
Lactate dehydrogenase level (ratio of upper normal reference)	116	1.03	0.04	565	1.14	0.03	0.019

* Two independent sample comparison between patients being heterozygous and homozygous for the least prevalent allele (G-carrier, GA/GG) versus patients being homozygous for the most prevalent allele (non G-carrier, AA)

Table 2. Frequencies of C-SIGN-336A/G genotypes and alleles in subgroups of SARS patients

Variable	Lactate dehydrogenase levels on admission*		P value	Odds ratio (95% CI)
Genotype	Low (n=585)	High (n=96)		
GG	3 (1%)	0 (0%)	-	-
GA	105 (18%)	8 (8%)	0.047	-
AA	477 (82%)	88 (92%)	-	-
Non G-carrier (AA)	477 (82%)	88 (92%)	Reference	
G-carrier (GA/GG)	108 (18%)	8 (8%)	0.014	0.40 (0.19-0.85)
Alleles	Low (n=1170)	High (n=192)		
A	1059 (91%)	184 (96%)	Reference	
G	111 (9%)	8 (4%)	0.015	0.41 (0.20-0.86)

* Cutoff value of normalised LDH level is 1.6 fold of upper normal reference

Table 3. χ^2 tests for lactate dehydrogenase (LDH) level and combined ICAM3/DC-SIGN genotyping of all patients

Combined genotype		LDH level*		P value†	Odd ratio (95% CI)
ICAM3 Asp143Gly (+443T/C)	DC-SIGN -336A/G	High (n=95)	Low (n=582)		
TT	AA	62 (65%)	377 (65%)	-	Reference
TT	AG	6 (6%)	89 (15%)	0.039	0.41 (0.17-0.98)
TT	GG	0 (0%)	3 (1%)	1.000	-
TC	AA	20 (21%)	90 (15%)	0.286	1.35 (0.78-2.35)
TC	AG	2 (2%)	15 (3%)	1.000	-
CC	AA	5 (5%)	7 (1%)	0.021	4.34 (1.34-14.12)
CC	AG	0 (0%)	1 (0%)	1.000	-

* Cutoff value of normalised LDH level is 1.6 fold of upper normal reference

† P=0.024 for overall combined genotype, df=6

-336A allele.[1] Our EMSA results showed that the G probe (lane 13) had stronger binding capacity to the nuclear extract compared to the A probe (lane 8) [Fig 1]. The binding capacity of the G probe was effectively competed by the Sp1 consensus probe (lane 17) and by the unlabelled cold G probe (lane 15). This suggested that the DC-SIGN -336A/G SNP affected the binding of Sp1 present in the nuclear extracts. Results from luciferase reporter assays on the DC-SIGN promoter constructs containing the -336A/G SNP showed that the G allele promoter construct had significantly lower promoter activity compared to the A allele promoter construct (-336G:-336A=3.8:1, P=0.011) [Fig 2].

Our EMSA and promoter activity results suggested that transcription factor binding sites, such as Sp1, may span

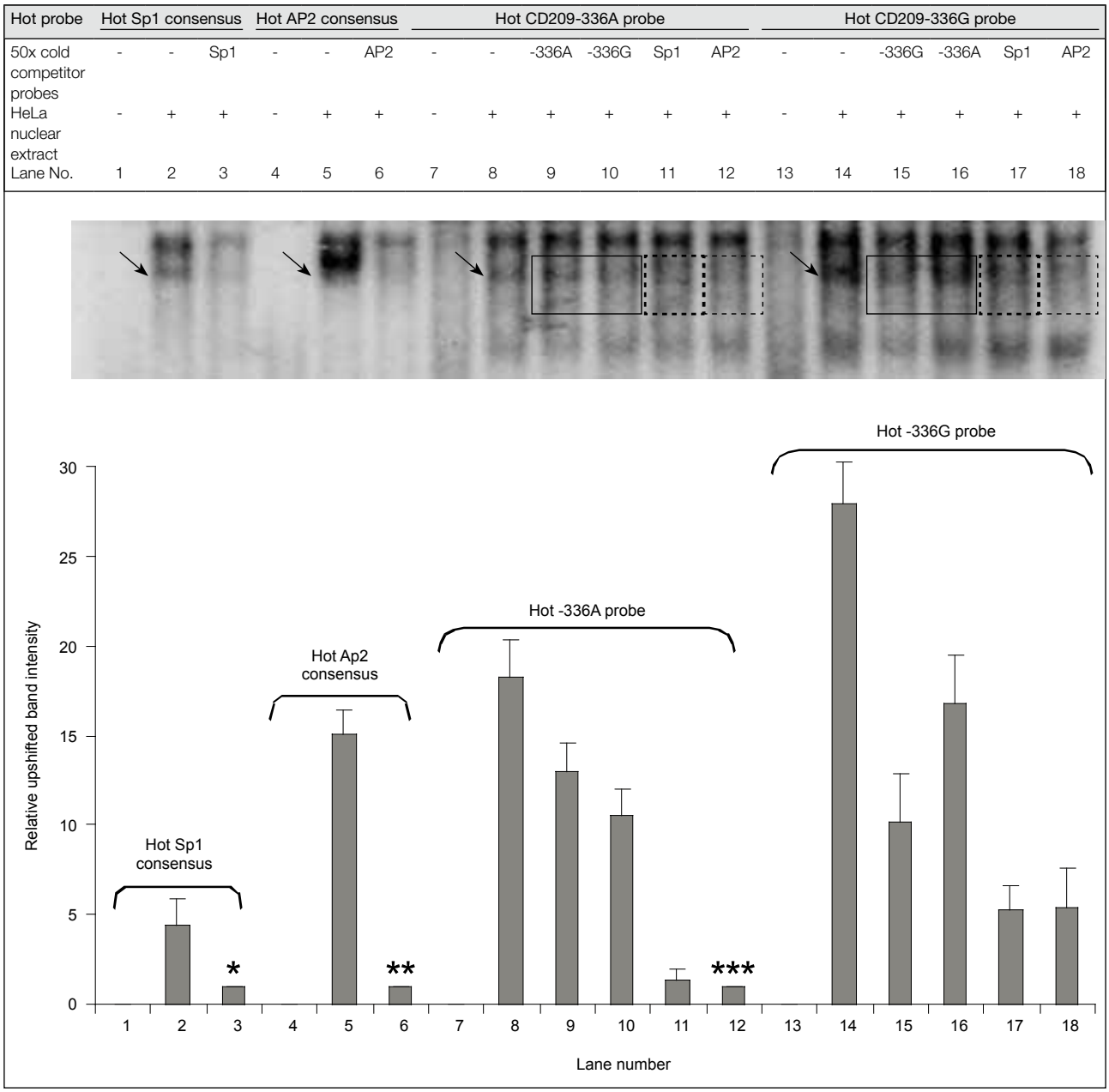


Fig 1. Electrophoretic mobility shift assay of Sp1 consensus, AP2 consensus, CD209 -336A and -336G probes using HeLa nuclear extract

The upshifted bands are indicated by arrows. The relative intensity of upshifted bands is plotted. The band intensity in lanes 3 and 6 are used to normalise against that in lanes 2 and 5 respectively, whereas the band intensity in lane 12 is used to normalise against that in lanes 8 to 11 and 13 to 18. Hence, Lanes 3, 6, and 12 are assigned an arbitrary relative value of 1. Bars denote standard deviations. The upshifted bands could be competed by respective cold competitor probes. The intensity of the upshifted band of DC-SIGN -336G probe (lane 14) is stronger than that of -336A probe (lane 8). The upshifted band of -336G probe is competed by cold -336G probe more effectively than by cold -336A probe (lanes 15 and 16), suggesting that -336G allele has stronger binding affinity to HeLa nuclear proteins than -336A allele

over the -336A/G SNP and hence modulate transcriptional activity. Our findings are in concordance with that reported by Sakuntabhai et al,[1] supporting the transcription regulatory role of the DC-SIGN -336A/G SNP.

DC-SIGN expression in PBMC derived DCs

On screening 20 consecutive PBMC samples, only two were found to carry the -336AG genotypes, whereas the rest were the homozygous -336AA genotype. The homozygous -336GG genotype was difficult to obtain as it contributed <1% in our population. As a result, two homozygous AA and two heterozygous AG genotypes PBMC samples were used for deriving DCs, which were subsequently used for detection of DC-SIGN expression by flow cytometry. Our results showed that the expression of DC-SIGN in PBMC derived DCs with the heterozygous -336AG genotype was not significantly different from that with -336AA genotype (P=0.767, *t*-test) [Fig 3]. However, the number of cases was small and did not allow conclusions to be made. Furthermore, the effect of genotype on surface expression of DC-SIGN may only be significant with the homozygous -366GG genotype, which we were unable to study owing to the rarity of this genotype.

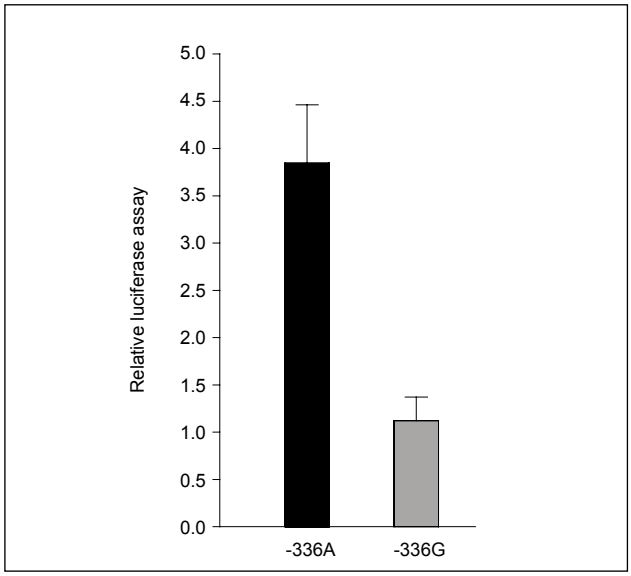


Fig 2. Luciferase assay results of the DC-SIGN -336A (pGL3-basic/DC-SIGN-336A) and -336G (pGL3-basic/DC-SIGN-336G) promoter constructs in HeLa cells

Results are the means from three independent triplicate experiments. The luciferase activity readings of -336A construct is normalised against that of -336G construct which is assigned an arbitrary relative value of 1. Luciferase levels for -336G are significantly lower (P=0.011, Student's *t* test with unequal variance). Error bars denote standard deviations

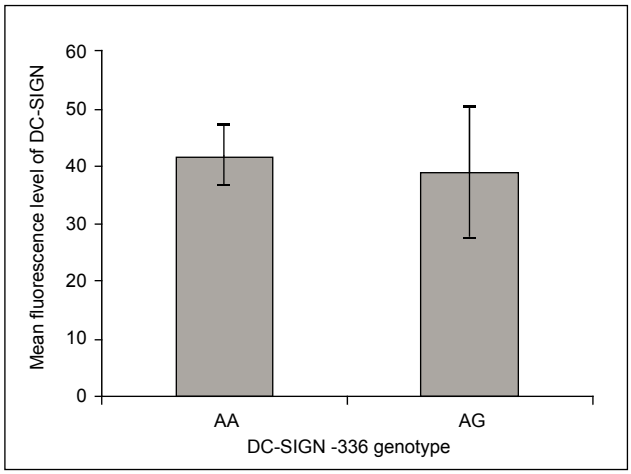


Fig 3. Comparison of mean fluorescence level of DC-SIGN of dendritic cells having -336AA or -336AG genotype

Difference in DC-SIGN level is not significant (P=0.767, Student's *t* test). Error bars denote standard deviations

Discussion

The G allele of the DC-SIGN -336A/G SNP has been shown to play a protective role against parenteral HIV-1 infection in European-American populations,[3] against dengue in Thailand,[1] and confer a low risk of tuberculosis in African populations.[2] Although the -336A/G SNP was not associated with susceptibility to SARS infection, the G allele of this SNP was associated with lower LDH levels in SARS patients being admitted to hospital apart from being independently associated with good prognosis for the disease.[4] The association remained significant after adjustment for patient age and gender in both overall cases and an unrelated patient subset.

Increase in serum LDH activity was postulated to be a result of massive tissue destruction during the acute phase of SARS-CoV infection. A high LDH level on admission was an independent prognostic indicator for severity of SARS infection and could help clinicians predict adverse clinical outcomes.[4] A high LDH level reflected tissue necrosis related to immune hyperactivity in SARS. A cytokine and chemokine storm occurs with a significant elevation of T-helper cell cytokine IFN-γ, inflammatory cytokines and Th1 IFN-γ-induced protein-10 during the early phase of SARS-CoV infection. The respiratory tracts of affected individuals who died during the first 10 days of illness showed diffuse alveolar damage with a mixed alveolar infiltrate, lung oedema and hyaline membrane formation indicating lung injury during progression of the disease.

Our EMSA and promoter activity results showed lower transcription activity of the -336G allele of DC-SIGN promoter compare to the -336A allele, suggesting that -336A/G has a functional role of in transcriptional regulation of DC-SIGN. Sp1 is a possible transcription factor binding site spanning the -336A/G SNP. Our findings are in concordance with those reported by Sakuntabhai et al[1] and suggest that individuals with the G allele might be less immune-responsive to SARS-CoV. Lower G-allele promoter activity may result in lower levels of DC-SIGN protein expressed, resulting in a lower T-cell response and reduced cytokine and chemokine secretion infiltrating the alveoli. Thus, the G allele of DC-SIGN -336A/G might

protect the lung from injury during the progression of SARS infection.

Interestingly, the frequency of the minor allele -336G identified varied in different ethnic groups; Asian populations have the lowest frequency. The frequency of the minor allele -336G in Chinese (8.7%) was similar to that in Thais (8%),[1] but was significantly lower than in European Caucasians, Africans, and American-Africans (>20%).[1,2] This genetic heterogeneity between different populations suggests that the DC-SIGN SNP might be associated with genetic predisposition to disease. Moreover, the significant association of the DC-SIGN SNP to various infectious diseases suggests that the genetic difference may be caused by natural selection of individuals who survived lethal disease/pathogen challenges.

Combined genotype analysis of ICAM3 +443T>C and DC-SIGN -336A>G showed that overall combined genotypes were significantly associated with disease severity. We previously reported that ICAM3 +443C carriers were associated with high LDH levels on admission. This study found that homozygous -336AA subjects were associated with high LDH levels on admission compared to -336G carriers. The combined +443CC/-336AA genotype was in agreement with this high LDH level association. Patients with this +443CC/-336AA genotype had a 4.3-fold greater risk than standardised LDH levels on admission. Nonetheless, multivariate logistic regression analysis was unable to demonstrate any synergistic effect between the two SNPs.

In summary, a SNP of DC-SIGN -336A/G is associated with LDH levels on admission, which is an independent prognostic indicator for the severity of SARS. This functional SNP affects the promoter activity of DC-SIGN and may alter gene expression and hence host immune response.

Acknowledgements

This study was supported by the Research Fund for the Control of Infectious Diseases, Food and Health Bureau, Hong Kong SAR Government (#04050252). We thank all colleagues, doctors, technical staff, patients, and participants who contributed to this project. The results of this study were published in:

Xu MS, Chan KY, Peiris JS, Yip SP, Cheung AN, Khoo US. A variant in the CD209 promoter is associated with severity of severe acute respiratory syndrome (SARS). The proceedings of the 11th HUGO's Human Genome Meeting, Helsinki, Finland, 2006.

Chan KY, Xu MS, Ching JC, et al. CD209 (DC-SIGN) -336A>G promoter polymorphism and severe acute respiratory syndrome in Hong Kong Chinese. Hum Immunol 2010;71:702-7.

References

1. Sakuntabhai A, Turbpaiboon C, Casademont I, et al. A variant in the CD209 promoter is associated with severity of dengue disease. Nat Genet 2005;37:507-13.
2. Barreiro LB, Neyrolles O, Babb CL, et al. Promoter variation in the DC-SIGN-encoding gene CD209 is associated with tuberculosis. PLoS Med 2006;3:e20.
3. Martin MP, Lederman MM, Hutcheson HB, et al. Association of DC-SIGN promoter polymorphism with increased risk for parenteral, but not mucosal, acquisition of human immunodeficiency virus type 1 infection. J Virol 2004;78:14053-6.
4. Chan KY, Ching JC, Xu MS, et al. Association of ICAM3 genetic variant with severe acute respiratory syndrome. J Infect Dis 2007;196:271-80.
5. Lee N, Hui D, Wu A, et al. A major outbreak of severe acute respiratory syndrome in Hong Kong. N Engl J Med 2003;348:1986-94.

M Small 司馬樂
CK Tse 謝智剛

Complex network models of disease propagation: modelling, predicting and assessing the transmission of SARS

Key Messages

1 Complex network models improve descriptions of disease transmission, and reveal the relationship and contacts between all individuals within a community.
2. These models might have been able to predict the extent of the SARS outbreak. For realistic parameter values (estimated from community demographics), the outbreak was substantial only if transmission within hospitals was not controlled.
3. Super-spreading events can occur as a consequence of human connectivity, and are not necessarily a result of highly infectious individuals.
4. Avian influenza can be described by the scale-free network, which exhibits occasional extremely high connectivity between outbreak sites. This means that merely reducing infectivity cannot effectively control avian influenza. Control can only be achieved by removing places where large numbers of poultry congregate.

Hong Kong Med J 2010;16(Suppl 4):S43-4

Department of Electronic and Information Engineering, Hong Kong Polytechnic University
M Small, CK Tse

RFCID project number: 03040102

Principal applicant and corresponding author:
Dr Michael Small
Department of Electronic and Information Engineering, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong SAR, China
Tel: (852) 2766 4744
Fax: (852) 2362 8439
Email: ensmall@polyu.edu.hk

Introduction

Traditional (compartmental) models of disease transmission categorise individuals from a population, based on their current pathology. These models trace the proportion of a population based on which state (susceptible, exposed, infected, or resistant to a disease) into which individuals fall. These models provide a population-based description that offers a stabilising response to the presence of an infectious agent. We examined data from two recent Hong Kong outbreaks, namely the current global avian influenza outbreak among birds and the 2003 outbreak of severe acute respiratory syndrome (SARS) among humans. Both cases exhibited large variability in terms of outbreak intensity and hence were not consistent with the standard models of disease transmission. The observed behaviour can be more readily explained using a computational (discrete agent-based) model of transmission on a complex networks. We described a model by which networks consistent with this data may arise in practice, and then provided a mathematical analysis of the epidemic dynamics and spreading behaviour.

Methods

This study was conducted from November 2005 to November 2007. The focus was on the application of computational complex network-based models to describe the SARS outbreak in Hong Kong. Standard models of disease transmission assume that all individuals have an equally small probability of being infected by an infectious person in the community. In contrast, a complex network model actually models the connectivity between individuals and only allows for the possibility of transmission of an infectious agent along those connections. A network consists of a set of nodes, with a set of edges connecting the nodes. Each node in the network corresponds to an individual within the community and a connection is drawn between two nodes (individuals) if there is a non-zero probability of infection. That is, two individuals are connected if they have contact that is sufficiently intimate to allow the disease to pass between them. For example, for a sexually transmitted disease this would require intercourse. For SARS (assuming droplet transmission), this only requires that the two individuals be in physical contact, or close proximity, for some period of time.

It is not feasible to know the full network of connections between individuals in a city of seven million people. Nonetheless, we know a lot about the general structure of that connectivity. Many studies of social phenomena have shown that the connectivity between individuals follows particular patterns. We implemented these patterns in our model. In particular, for human social groups, the individuals typically form a small-world network. That is, one's neighbours on a network typically are neighbours themselves (that is, your friends are usually friends with each other too). However, at the same time each node has some random connections, which violate this pattern. It has been shown that under these conditions the average path between individuals (friend of a friend of a friend and so on) typically consists of a small number of connections. A short chain of mutual acquaintances connects random pairs of individuals. This effect is known as the 'six degrees of separation' and has been observed in diverse social systems: from Hollywood movies, to co-

authorship of condensed matter physics manuscripts, to internet and email connectivity. The same is also true for social and collegial groupings within a city. We applied this model to the population of Hong Kong and restricted the transmission of a virus to pathways in this network.

Results

Using complex network models (and suitable parameter values chosen from measurable demographic features of the community), computer simulation of SARS transmission matched qualitatively and quantitatively the data observed in Hong Kong. In particular, the clustering of the outbreak in Amoy Gardens and the Prince of Wales Hospital emerged as a natural feature of the model structure. Moreover, the random variability in the degree of connectivity allowed for super-spreader events (localised highly effective transmission) without a need to explicitly include this feature in the model of infectivity. This is a significant improvement on existing models. These complex network models provided a simpler and more effective description of reality, although individual events, such as those cited above, had specific detailed causes. Our models also showed that the degree of infection in the 2003 outbreak in Hong Kong is due primarily to transmission within hospitals. With proper infection control within hospitals, the outbreak would probably have been far less severe.

Nonetheless, given that hospital-based infection did occur (at least initially), our simulations provided an ensemble of predicted outcomes. The actual outcome was typical of that ensemble, but the range of severity was wide. We observed typical outbreaks with fewer than 10 to more than 10 000 casualties.

Discussion

Our simulations for SARS showed a good agreement with reality, but the model itself is not derived from reality. To obtain the actual infection chains between individuals, we studied an outbreak of avian influenza, particularly for animal-to-animal infection, as data pertaining to times and locations of such an outbreak are widely available.

We used these data to construct a network of potential connection pathways to test the hypothesis that this network is complex and pertaining to a small world (consistent with that for SARS in Hong Kong). We found that the network obtained in this way is complex, but not of a small world. The reason for this is a consequence of the global geographical distribution of outbreak sites; geographically distant sites cannot be connected through a short path. Nonetheless, we discovered another important property in this network. We found that the network of avian influenza infection pathways is scale-free. This means that, on average, the number of secondary infection sites connected to a primary site is very large (it follows a power-law distribution). Although most sites are connected to only a small number of other sites, some sites have a very large number of connections. Moreover, for the parameter values observed, the best guess (on average) that one can make for the number of connections from a given site is infinite (and yet most sites have very few connections). Networks such as this have been theorised by physicists to model disease transmission; we demonstrated evidence of such a network occurring in nature.

The theory presented in the physics community implies that for disease transmission on such a network the disease will continue to exist and be transmitted provided the rate at which infection passes between individuals is greater than zero. This is in sharp contrast to standard epidemiological models where a disease can be eliminated by reducing the infectivity below some finite threshold. For avian influenza this is not possible. The only way in which avian influenza can be eliminated is by changing the structural features of the network of infection pathways. The practical implication of this is that all places where a large number of live poultry are brought together must be eliminated (or at least very strictly controlled). The only other successful control measure is a thorough cull, as implemented in Hong Kong in 1997.

In addition to the data driven analysis described above, we conducted an extensive new mathematical analysis on the transmission of infectious agents on networks. We studied more physically realistic models of disease transmission and examined under which circumstances non-zero thresholds would emerge. The threshold is the critical value of infectivity or transmissibility of a disease. If the rate of infection exceeds this value, then the disease will become endemic, otherwise the disease will be self-terminating. For diseases transmitting on a general class of scale-free networks (including that for avian influenza), it has been shown that the threshold must be zero. Our theoretical work aimed to address the following question: which modifications to this network allow for a non-zero threshold? Such modifications then allow us to plan appropriate control strategies. Our work also examined models for the various control measures that may be implemented, and efforts to economically modify the network topology to allow for easier control of disease transmission.

All the models described in this work, together with the publicly available data are provided as a self-contained software package from the website: http://small.eie.polyu.edu.hk/jema/

Acknowledgement

This study was supported by the Research Fund for Control of Infectious Diseases, Food and Health Bureau, Hong Kong SAR Government (#03040102).

KY Yuen 袁國勇
PCY Woo 胡釗逸
SKP Lau 劉嘉珮

A multilocus sequence typing system for *Penicillium marneffei:* an international molecular cyber system for tracking its origin and transmission

Key Messages

1. A multilocus sequence typing (MLST) system was developed for *Penicillium marneffei*, an emerging opportunistic infection in Southeast Asia.
2. A website for *P marneffei* MLST is available at http://mlstdb.hku.hk:14206/MLST_index.html to enable comparison of strains from different localities.
3. Our *P marneffei* MLST system appears to be more discriminating and more suitable for epidemiology studies than other similar systems.
4. Non-housekeeping genes should be incorporated into the MLST system to achieve greater discriminatory power.

Introduction

Penicillium marneffei is the most important dimorphic fungus causing systemic mycosis in Southeast Asia.[1] The global dissemination of the human immunodeficiency virus (HIV) has led to the emergence of this infection as an important opportunistic mycosis in HIV-positive patients. About 10% of acquired immunodeficiency syndrome patients in Hong Kong are infected with *P marneffei*.

A highly reproducible and discriminative typing system enables better understanding of the epidemiology of the fungus. Although multilocus microsatellite typing and pulsed field gel electrophoresis were used to develop typing schemes for *P marneffei*, the results are not suitable for comparison due to experimental variations among different laboratories. We describe a multilocus sequence typing (MLST) scheme of *P marneffei*.

Methods

This study was conducted from January 2006 to December 2007. In order to have a better understanding of the epidemiology of *P marneffei*, we developed an MLST system for *P marneffei* using 44 human isolates.[2] The nucleotide sequences of the 11 housekeeping genes and 14 mannoprotein superfamily homologues in *P marneffei* strains were aligned and compared. The 44 *P marneffei* isolates were sequenced with the five most discriminative loci.

In the first part, DNA extracted from 10 of the 44 strains of *P marneffei* was used as the template for amplification of 11 housekeeping genes (mannose phosphate isomerase, plasma membrane H+ ATPase, pyruvate kinase, glutamate dehydrogenase, phosphoglucomutase, ribonucleoside-diphosphate reductase, glutamate synthase precursor, ribonucleotide reductase, transcription factor PacC, carbon catabolic repressor protein, and DNA topoisomerase II) and 14 lineage-specific genes (including MP1 and its homologues that belong to a novel mannoprotein superfamily).[2]

In the second part, DNA extracted from all 44 strains of *P marneffei* was used as the template for amplification of five of the 14 gene loci of the MP1 homologues (*MP1*, *MPLP4*, *MPLP7*, *MPLP10* and *MPLP13*).[2] The primers were designed by multiple alignments of the 14 homologues so that they were specific to each *MP1* homologue. The nucleotide sequences of the five gene loci in all the *P marneffei* strains were aligned and compared with those of strain PM1 using ClustalX (1.83). The ratio of non-synonymous to synonymous base substitutions (dn/ds) was calculated with START2 (http://pubmlst.org/software/analysis). Construction of dendrograms was performed with the unweighted pair group method with arithmetic mean (UPGMA) using Molecular Evolutionary Genetics Analysis 3.1. Grouping of sequence types (STs) into lineages was performed with BURST (Based Upon Related Sequence Types, http://pubmlst.

Hong Kong Med J 2010;16(Suppl 4):S45-6

Department of Microbiology, The University of Hong Kong
KY Yuen, PCY Woo, SKP Lau

RFCID project number: 03040282

Principal applicant and corresponding author:
Prof Kwok-yung Yuen
Department of Microbiology, The University of Hong Kong, University Pathology Building, Queen Mary Hospital, Pokfulam, Hong Kong SAR, China
Tel: (852) 2855 4892
Fax: (852) 2855 1241
Email: kyyuen@hkucc.hku.hk

org/analysis/burst/burst.shtml).

Results

Amplification and sequencing of the 11 housekeeping genes from 10 strains of *P marneffei* showed that the nucleotide sequences of all 11 genes were identical.[2]

In the first 10 strains of *P marneffei* sequenced, the sequences of four (*MPLP2*, *MPLP3*, *MPLP6* and *MPLP12*) of the 14 homologues were identical, five (*MP1*, *MPLP4*, *MPLP7*, *MPLP10* and *MPLP13*) showed more variations than others. Therefore, only these five loci were sequenced for the other 34 strains of *P marneffei* and the sequences of these five loci were used for developing the MLST system.[2]

Among the 2201 bp of the five loci, a total of 183 polymorphic sites were observed in the 44 *P marneffei* strains.[2] Allelic profiles were assigned to the 44 *P marneffei* strains. The alleles defined for the MLST system were based on sequence lengths of between 337 bp (*MPLP4*) and 549 bp (*MP1*).[2] The relatively high d_n/d_s ratio for the five genes indicated that a strong positive selective pressure was present at these loci.[2]

A total of 35 different STs were assigned to the 44 *P marneffei* isolates.[2] The overall discriminatory power for the MLST scheme was 0.9884.[2] The UPGMA was used to construct a dendrogram using the concatenated nucleotide sequences of the five gene loci from the 44 isolates.[2] The isolates were grouped into four lineages by BURST, with six STs in group 1, five in group 2, six in group 3 and two in group 4, whereas 16 STs did not belong to any of the four groups.[2]

Discussion

We developed an MLST system for *P marneffei*. The sequences of MP1 homologues are more variable than those of its housekeeping genes. The cloning and characterisation of MP1, which encodes an abundant, secreted and cell wall immunogenic mannoprotein, Mp1p, from *P marneffei* was reported in 1998.[3] Genes with higher lineage specificity in ascomycetes evolve at a much faster rate than those with lower lineage specificity.[4] Therefore, genes with higher lineage specificity are potentially more useful targets than housekeeping genes for typing pathogenic fungi. Furthermore, Mp1p is a cell wall immunogenic protein located on the surface of *P marneffei*, and hence is subject to strong selective pressure by the immune system.[3] Therefore, MP1 homologues are potentially more rapidly evolving than housekeeping genes. According to our complete genome sequence project,[5] there are more than 10 *MP1* homologues in the *P marneffei* genome. We therefore hypothesised that it may be more discriminatory if this set of gene targets were used for MLST in *P marneffei*. We sequenced 11 housekeeping genes and the *MP1* homologues of *P marneffei*. The sequences of the 11 housekeeping genes were identical among the 10 strains of *P marneffei* sequenced, but remarkable variations exist in the sequences of the *MP1* homologues. Therefore, the *MP1* homologues were used as targets for building the MLST system in *P marneffei*.

Our study showed that lineage specific genes may be better candidates than housekeeping genes for sequence based typing. When MLST systems were used in fungal pathogens, the discriminatory power was often unsatisfactory whenever only housekeeping genes were used. When one or two non-housekeeping genes were also included, the discriminatory power was markedly improved. This is in concordance with our previous observation that lineage specific genes were associated with more rapid evolutionary rates.[4] Therefore, lineage specific genes could be better targets for MLST schemes when applied to slowly evolving or recently evolved pathogens.

Acknowledgement

This study was supported by the Research Fund for the Control of Infectious Diseases, Food and Health Bureau, Hong Kong SAR Government (#03040282).

References

1. Yuen KY, Wong SS, Tsang DN, Chau PY. Serodiagnosis of P*enicillium marneffei* infection. Lancet 1994;344:444-5.
2. Woo PC, Lau CC, Chong KT, et al. MP1 homologue-based multilocus sequence system for typing the pathogenic fungus *Penicillium marneffei*: a novel approach using lineage-specific genes. J Clin Microbiol 2007;45:3647-54.
3. Cao L, Chan CM, Lee C, Wong SS, Yuen KY. MP1 encodes an abundant and highly antigenic cell wall mannoprotein in the pathogenic fungus *Penicillium marneffei*. Infect Immun 1998;66:966-73.
4. Cai JJ, Woo PC, Lau SK, Smith DK, Yuen KY. Accelerated evolutionary rate may be responsible for the emergence of lineage-specific genes in ascomycota. J Mol Evol 2006;63:1-11.
5. Woo PC, Zhen H, Cai JJ, et al. The mitochondrial genome of the thermal dimorphic fungus *Penicillium marneffei* is more closely related to those of molds than yeasts. FEBS Lett 2003;555:469-77.

AUTHOR INDEX

Published by the Hong Kong Academy of Medicine Press for the Government of the Hong Kong Special Administrative Region. The opinions expressed in the *Hong Kong Medical Journal* and its supplements are those of the authors and do not reflect the official policies of the Hong Kong Academy of Medicine, the Hong Kong Medical Association, the institutions to which the authors are affiliated, or those of the publisher.