Karolinska Institutet

Institutionen för Medicinsk Epidemiologi och Biostatistik

Molecular Epidemiologic Studies on *Helicobacter pylori* Infection and Stomach Cancer Risk

AKADEMISK AVHANDLING
som för avläggande av medicine doktorsexamen vid Karolinska Institutet offentligen försvaras i Petrénsalen, Nobels väg 12B.

Tisdagen den 23. augusti 2011, kl 9:00

av

Zongli Zheng

Huvudhandledare:
Professor Weimin Ye
Institutionen för Medicinsk Epidemiologi och Biostatistik
Karolinska Institutet

Bihandledare:
Professor Olof Nyrén
Institutionen för Medicinsk Epidemiologi och Biostatistik
Karolinska Institutet

Professor Lars Engstrand
Institutionen för Mikrobiologi, Tumör- och Cellbiologi
Karolinska Institutet

Fakultetsopponent:
Professor Ernst J. Kuipers
Dept. of Gastroenterology and Hepatology
Erasmus University

Betygsnämnd:
Professor Gunnar Steineck
Institutionen för Onkologi och Patologi
Enheten för Klinisk Cancerepidemiologi
Karolinska Universitetssjukhuset

Professor Mats Nilsson
Institutionen för Immunologi, Genetik och Patologi
Uppsala Universitet

Professor Ann-Mari Svennerholm
Avd för Mikrobiologi och Immunologi
Göteborgs Universitet

Stockholm 2011

ABSTRACT

Helicobacter pylori (H. pylori) infection increases stomach cancer risk. The aim of this thesis was to study genetic susceptibility from the host and to develop molecular methods for future characterization of bacterial virulence factors in longitudinal cohorts.

In Study I, we investigated the association between genetic variation in an O-glycan transferase encoding gene (*a4GnT)* and *H. pylori* infection and gastric cancer risk in a Polish population-based case–control study (273 gastric cancer patients and 377 controls). A haplotype at rs2622694–rs397266 was associated with *H. pylori* infection, with the A-A haplotype associated with a higher risk compared with the most frequent G-G haplotype (odds ratio 2.30; 95% confidence intervals 1.35–3.92). Neither this haplotype nor the tagSNPs were associated with overall gastric cancer risk.

In Study II, we characterized genomic evolution of *H. pylori* over 20 years in the stomach. Whole genome of 21 sequential isolates 20 years apart, from 7 patients, were sequenced using 454 sequencing platform. There were on average 260 point mutations (range 70 to 488) per isolate over 20 years, and 45 recombinations (range 18 to 92). Genes in the cell motility category were overrepresented in point mutations and recombinations. Specifically, mutations often affected genes involved in chemotaxis, vacuolating cytotoxin-like protein, restriction and type IV secretory pathway; and recombinations affected glycosyltransferase involved in lipopolysaccharide biosynthesis. The major form of single nucleotide substitutions was transition (85%) and the minor form was transversion (15%). Mutation was sequence context-dependent.

Clinical samples are often precious and of trace amounts. In Study III, we developed novel methods for DNA shotgun library construction and quantification. As compared with the standard procedure, our double-stranded and Y library construction methods are simpler and more efficient. A highly sensitive Taqman MGB-probe-based quantitative polymerase chain reaction (qPCR) was developed to quantify the amount of effective library. We also demonstrated that the distribution of library molecules on capture beads follows a Poisson distribution. Combining the qPCR and Poisson statistics, the labor intensive and costly titration can be eliminated and trace amounts of starting material is applicable.

Archived formalin-fixed and paraffin-embedded (FFPE) biopsies, coupled with long term follow-up, are valuable resources for molecular epidemiologic studies. Study IV presented a method based on laser capture micro-dissection and modified whole genome sequencing methods to obtain metagenomic profiles of *H. pylori* from 15-year old FFPE biopsy sections.

Keywords: *Helicobacter pylori,* Stomach cancer, Next-generation sequencing, Formalin-fixed and paraffin-embedded biopsy