Prof. Dr. Jean-Claude Piffaretti
Istituto Cantonale
di Microbiologia
Via Mirasole 22A
6501 Bellinzona
piffaretti@interlifescience.ch

4049-063270
01.09.2001-31.12.2005

Molecular evolution of pathogenicity and resistance in Escherichia coli

Objectives The aim of this project was to analyse how genes encoding antibiotic resistance and virulence determinants evolve in *Escherichia coli*, a bacteria that is both commensal and pathogenic in humans and animals. Under investigation were both food animals (cattle, pigs and poultry) and pets (cats and dogs).

Conclusions Through comparative genome hybridisation (CGH) we could confirm that the *E. coli* genome is highly plastic, with a bipartition into a stable (constant) core gene pool containing basic housekeeping genes and a flexible (divergent) gene pool that harbours pathogenicity and resistance genes. Differences in gene content allowed the separation of the *E. coli* strains according to their host origin. Strains isolated from pets correlated with phylogenetic group B2 (group B2 together with group D includes nearly all extra-intestinal disease strains). Forty percent of the analysed *E. coli* strains were shown to harbour integrons that were found to be widely distributed in both human clinical and animal strains but were more prevalent in strains belonging to phylogenetic group A.
Furthermore, using DNA array and CGH analysis, 105 *E. coli* strains isolated from different hosts and pathologies were tested on microarrays including *E. coli* K12, CFT073 (uropathogen isolate) and enterohemorrhagic O157:H7 specific oligonucleotides along with 15 virulence genes. The results clearly indicate that the genetic diversity among different *E. coli* isolates is very high, with evidence of genetic exchanges. Certain virulence genes were shown to be found more often outside group B2, i.e. in phylogenetic group A (which includes most commensal and diarrhoeagenic strains), B1 or D, or are broadly distributed without focal presence in any of these groups.
To conclude, the correlation observed between integron carriage and food animals suggests that animals raised for economic purposes might be subjected to a major antibiotic pressure. Given the evidence of horizontal gene transfer in *E. coli*, healthy animals may represent an important reservoir of genes (which may contribute to *E. coli* pathogenicity) that might be transferred to other microorganisms by horizontal gene transfer. Finally, for the species *E. coli* the concepts “virulence”, “fitness” and “colonisation factor” seem to be overlapping, and we are still far from understanding the relationship between virulence and commensalisms in *E. coli*.

Main results and findings

Microarray development and use for phylogenetic studies The oligonucleotides were designed using *E. coli* K12 genome sequence database, an appropriate BLAST program and a set of 2,700 specific oligonucleotides representing different genes from *E. coli* K12. Comparison of microarrays data with sequence data (MLST) for 8 housekeeping genes (*pgi, icd, arcA, aroE, rpoS, mdh, mtlD, gyrB*) for 19 *E. coli* strains revealed that the two dendrograms were not coherent. This is probably due to the fact that MLST is based on the variability of housekeeping genes (in principle selectivity neutral, hence rather stable), whereas microarrays are based on the presence or divergence of variable genes.

***E. coli* genome comparison of human and animal commensal isolates using microarrays** The comparative genome hybridisation (CGH) analysis of 19 commensal *E. coli* strains collected from the intestinal flora of human and diverse animals by microarrays revealed or confirmed the following:

- The *E. coli* genome is highly plastic, with a bipartition into a stable (constant) core gene pool containing basic housekeeping genes and a flexible (divergent) gene pool.
- The divergent gene pool was further subdivided into moderately divergent (MD) and hyperdivergent (HD) gene groups.
- HD genes show a chromosomal clustering and are characterised by a different GC content than the remaining of the *E. coli* genome. These observations corroborate the hypothesis that HD genes arose by horizontal gene transfer.

- The majority of divergent coding sequences are found in "cellular processes" (i.e. defence mechanisms, cell motility, intracellular trafficking) and harbour pathogenicity and resistance genes. It is to be noted that an incremented plasticity of these genes may favour adaptability to highly variable environments.
- The constant coding sequences are more frequently associated to the "information storage and processing".
- A significantly higher proportion of divergent genes is observed in humans and pets than in strains isolated from food animals.
- Strains isolated from pets correlate with phylogenetic group B2 (group B2 together with group D includes nearly all extra-intestinal disease strains).

Distribution and characterisation of integrons in 120 *E. coli* strains of human and animal origin
Forty percent of the analysed *E. coli* strains were shown to harbour integrons. However, none of the strains were found to contain both class 1 and 2 integrons. There was a significant statistical correlation between class 1 integron carriage and the three host categories (food animals, pets and humans). Integrons were more prevalent in phylogenetic group A strains (which include most commensal and diarrhoeagenic strains) than in groups B1, B2 and D.

Comparative genome hybridisation of 105 *E. coli* strains isolated from different hosts and pathologies by means of DNA array and CGH analysis. The microarrays used for this study included *E. coli* K12, CFT073 (uropathogen isolate) and enterohemorrhagic O157:H7 specific oligonucleotides along with 15 virulence genes. The results can be summarised as follows:

- There is a separation of the invariant component of the genome (i.e. the core genome) from the remaining genes (flexible gene pool).
- There is a high level of genetic diversity among different *E. coli* isolates, and there is evidence of genetic exchanges.
- Certain virulence genes (i.e. operons *ddp, foc, hyf, Isr, paa, sfm, ycb* and *yra*) are found more often outside group B2, i.e. in phylogenetic group A, B1 or D, or are broadly distributed without focal presence in any of these groups.
- There was no association between genome structure and origin of the *E. coli* strains (host and pathology). However, an association between absence of specific genes and affiliation to a particular phylogenetic group (i.e. group B2 and/or D) could be observed.
- There was no association between the presence or divergence of virulent determinants and particular disease phenotype.

Publications of the NRP 49 project

Grasselli E, François P, Gutacker M, Gettler B, Benagli C, Convert M, Boerlin P, Schrenzel J, Piffaretti J-C.
Genome comparison of human and animal commensal *Escherichia coli* strains by array typing: evidence for horizontal gene transfer.
Submitted for publication.

Cocchi S, Grasselli E, Gutacker M, Benagli C, Convert M, Piffaretti J-C.
Distribution and characterisation of integrons in *Escherichia coli* strains of animal and human origin.
FEMS Immunol. & Medical Microbiol.: In press 2007.

back to the NRP 49 final report