F. Rodríguez-Valera[1*], J. García-Martínez[1], A. Zorraquino[2] and A.J. Martínez-Murcial

Genetic Diversity in Uropathogenic *Escherichia coli* as Shown by RAPD

(Random Amplified Polymorphic DNA), 16S-23S rDNA spacers and adhesin gene restriction patterns

Abstract

A collection of *Escherichia coli* strains isolated from urinary tract infections (UTI) and some representatives of the ECOR collection were characterized by using three different types of molecular markers. The strains were typed by RAPD (Random Amplified Polymorphic DNA). The 16S rDNA-23S rDNA spacers were amplified by PCR and subjected to restriction analysis. The presence or absence of G adhesins type II and/or III have also been assayed. Analysis of the spacers restriction patterns showed two markedly differentiated clusters designated α and β. Both RAPD and spacer restriction patterns originated similar clusters of strains showing a consistency in the evolution of the global genome with the sequence variation of the ribosomal spacers. Furthermore, the G adhesin type (and involvement in upper urinary tract infection) correlated well with the spacer groupings. The α and β clusters could be intraspecific groups produced by partial sexual isolation or other barriers that are originating a divergent evolution.

Molecular methods allow an extremely fine strain typing that can be used to establish the population structure of bacterial species. Here we have used three molecular approaches to characterize a collection of uropathogenic *Escherichia coli* (74 strains) obtained from three hospitals located in geographically distant towns in Spain, six representatives of the ECOR collection and other reference strains. The three markers investigated represent different levels of extension and, theoretically at least, of conservation of the genome. RAPD (Random Amplified Polymorphic DNA) can characterize a bacterial strain to the level of defining individual clones (only individuals sharing a very close common ancestor show identical RAPD patterns). To study intraspecies genomic diver-

[1] Departamento de Genética y Microbiología, Universidad de Alicante, Campus de San Juan, Apartado 374, 03080 Alicante, Spain.
[2] Hospital Universitario San Juan de Alicante, Carretera de Valencia s/n, San Juan de Alicante, Alicante, Spain.

sity, a RAPD band (characterized by size, but if required it can be hybridized or sequenced) can be used as a genetic marker, the absence of the band in one strain is a polymorphism affecting this locus (a different allele). If the band is present in two strains they share the same allele at this locus (Tibayrenc *et al.*, 1993). The amount of information obtained in a short time is enormous since an unlimited number of primers can be used and it requires very little time or investment. In addition RAPD allows a typing of the total genome with no distinction of the gene type or sequence.

The 16S rDNA-23S rDNA spacers were amplified by PCR and subjected to restriction analysis. Parts of them are stretches of non-coding DNA, presumably non-functional, and should exhibit a considerable degree of sequence variation by genetic drift. These sequence polymorphisms could make the spacer a fast molecular chronometer to measure short term phylogeny, i.e. a good marker of major intraspecies lineages. Recently, length and sequence variation of 16S-23S rRNA spacers have been used to discriminate between different species and strains of bacteria (Barry *et al.*, 1991; Jensen *et al.*, 1993, Kostman *et al.*, 1992; Dolzani *et al.*, 1994).

Finally, the presence and type of G adhesins in *Escherichia coli* were determined by PCR amplification followed by digestion with restriction enzymes. The presence of adhesins and P fimbriae are virulence factors that facilitate the colonization by *E. coli* of the upper urinary tract (Ikäheimo *et al.*, 1994). Three types (I, II and III) are known, with different binding specificities (Hultgren *et al.*, 1993). The possession of one type or other of adhesin genes has been considered as a trait that can be easily won or lost and horizontally transferred (Plos *et al.*, 1989).

All 16S-23S spacers amplified with 16S14F-231R were digested with endonucleases *Alu*I, *Cfo*I, *Dde*I, *Hae*III, *Hinf*I, *Rsa*I and *Taq*I, but polymorphisms were only obtained with *Hinf*I, *Rsa*I and *Taq*I, which were selected for comprehensive analysis. Twenty-six different patterns were distinguished from the 16S-23S spacer RFLPs. The data were analyzed by calculating the Jaccard coefficient of each pairwise comparison, the profile relationships were analyzed by generating the UPGMA dendrogram. The strains clustered in two large groups, named Groups α and β, with similarity levels below 70%. Group α contained 48 strains including ATCC25922, ECOR35, ECOR52 and ECOR58. Twenty-four (50%) of the strains within this group had identical RFLP patterns, conforming a sub-group, $\alpha 2$, represented by ECOR52. Group β comprised 33 strains including K12 (CECT102), DHSα, ECOR10, ECOR49 and ECOR44. Strains K12 and ECOR10 had identical patterns. Strains in Group α are frequently negative for sucrose (81%) and raffinose (95%), while those in Group β are positive (60% and 52% respectively).

Twelve primers out of 28 assayed generated informative RAPD prints (ranging from 200 to 2700 bp) that were reproducible and they were selected. As expected a wide diversity was shown by RAPD and only in the case of strains isolated from the same individual, an obvious case of relapse, were identical patterns found. A total of 107 RAPD prints obtained with eleven primers were used

as markers for pattern comparison, and UPGMA relationships were calculated. The dendrogram of RAPD showed that most of the strains clustered with less than 80% similarity level which can be interpreted as having less than 80% common alleles for the loci studied. In spite of this high diversity two groups appear clearly separated with similarities around 40%. These two RAPD clusters precisely fit with the two groups (α and β) detected by RFLP of the spacers: 96% of the strains of one RAPD cluster belong to Group α and 90% of the other cluster to Group β of the spacer RFLP. Clustering by the UPGMA method has been carried out using the RAPD-prints of different primer combinations: primers A1, A2, A7, A8, A10 and A12; primers A16, A18, A28 and A29; and primers A2, A8, A12, A18 and A28. The clustering was well consistent in the four dendrograms obtained with different RAPD-print combinations and fits with Groups α and β at a similar level. It is noteworthy that reference strains clustered together in the same group in both spacer RFLP and RAPD dendrograms: ATCC25922, ECOR52, ECOR35 and ECOR58 in Group α and strains DH5α, R12 (CECT102), ECOR10, ECOR49 and ECOR44 in Group β. The fact that both RAPD and spacer restriction patterns originated similar clusters of strains showed a consistency of the evolution of the global genome and the sequence variation of the ribosomal spacers. The stability of the two groups, regardless of the number/type of genetic markers or group of strains used, is very high indicating that this grouping actually reflects a natural subdivision of the species.

A total of 32 strains contained one or more copies of the G adhesin gene, 21 of which were class II, 4 were class III and 7 strains contained both classes. The presence of adhesins was widely distributed among the strains of Group α: 50% of strains within this group showed patterns corresponding to classes II or/and III. Only 18% of Group β strains contained G adhesin Class II (6 out of 32) and none had type III. Although it is known that these genes are horizontally transferred (Plos *et al.*, 1989), the presence of G-adhesin was found to be more associated to Group α of UTI *E. coli* strains. Also the involvement in upper urinary tract infection correlated well with the spacer polymorphism, most of the strains isolated from pyelonephritis (6 out of 7) corresponded to the α rRNA spacer Group. This illustrates that in spite of the proven transferability of the adhesin genes, this happens rarely between strains of the α and β groups and this could also be true for other horizontally transferable markers.

The differences found between α and β Groups could be explained by a barrier to horizontal genetic transfer between the two groups, for example, due to a different restriction modification system, sensitivity to phages or recognition by sexual fimbriae. In fact, α and β groups could well be populations on their way to becoming new species, or they must be considered as such, if the establishment of recombinative barriers is part of the speciation process in bacteria as has been suggested (Dykhuizen *et al.*, 1991). The α and β clusters could be intraspecific groups produced by partial sexual isolation or other barriers that are originating a divergent evolution.

References

Barry, T., G. Colleran, M. Glennon, L.K. Dunican and F. Gannon, 1991 - The 16S/23S ribosomal spacer region as a target for DNA probes to identify eubacteria. *PCR Meth. Appl.* **1**, 51–56.

Dolzani, L., E. Tonin, C. Lagatolla and C. Monti-Bragadin, 1994 - Typing of *Staphylococcus aureus* by amplification of the 16S-23S rRNA intergenic spacer sequences. *FEMS Microb. Lett.* **119**, 167–174.

Dykhuizen, D.E. and L. Green, 1991 - Recombination in *Escherichia coli* and the definition of biological species. *J. Bacteriol.* **173**, 7257–7268.

Hultgren, S.J., F. Jacob-Dubuisson, C.H. Jones and C.-I. Bränden, 1993 - PapD and superfamily of periplasmic immunoglobulin-like pilus chaperones. *Adv. Protein Chem.* **44**, 99–123.

Ikäheimo, R., A. Siitonen, U. Karkkäinen, J. Mustonen, T. Heiskanen and P.H. Mäkela, 1994 - Community-acquired pyelonephritis in adults: characteristics of *E. coli* isolates in bacteremic and non-bacteremic patients. *Scandinavian J. Infect. Dis.* **26**, 289–296.

Jensen, M.A., J.A. Webster and N. Straus, 1993 - Rapid identification of bacteria on the basis of polymerase chain reaction-amplified ribosomal DNA spacer polymorphisms. *Appl. Environ. Microbiol.* **59**, 945–952.

Kostman, J.R., T.D. Edlind, J.J. LiPuma and T.L. Stull, 1992 - Molecular epidemiology of *Pseudomonas cepacia* determined by polymerase chain reaction ribotyping. *J. Clin. Microbiol.* **30**, 2084–2087.

Plos, K., S.I. Hull, R.A. Hull, B.R. Levin, F. Orskov and C. Suanborg-Eden, 1989 - Distribution of the P-associated-pilus (pap) region among *Escherichia coli* from natural sources: evidence for horizontal gene transfer. *Infect. Immun.* **57**, 1604–1611.

Tibayrenc, M., K. Neubauer, C. Barnabé, F. Guerrini, D. Skarecky and F.J. Ayala, 1993 - Genetic characterization of six parasitic protozoa: Parity between random-primer DNA typing and multilocus enzyme electrophoresis. *Proc. Nat. Acad. Sci., USA* **90**, 1335–1339.