3

Population Genetics of the "*Aeromonas hydrophila* Species Complex"

Mª Carmen Fusté, Maribel Farfán, David Miñana-Galbis,
Vicenta Albarral, Ariadna Sanglas and José Gaspar Lorén
Department of Health Microbiology and Parasitology, Faculty of Pharmacy,
University of Barcelona, Barcelona,
Spain

1. Introduction

Population genetics studies the genetic variability of individuals in a population based on the allele frequencies at several genes or loci and tries to explain this variability in terms of mutation, selection or genetic recombination. The statistical analysis of these frequencies allows models of evolution to be established, which will help us to understand and predict the past and present gene flow in the population (Maynard-Smith, 1991). For the most part population genetics has been designed for diploid organisms with sexual reproduction. In the words of Bruce Levin, "the genetic theory of adaptive evolution was developed by sexually reproducing eukaryotes, for sexually reproducing eukaryotes" (Levin & Bergstrom, 2000). As a consequence, before being applied to prokaryotes, population genetics needs to be adapted.

In theory the haploid nature of bacteria should simplify their analysis, since dominance or over-dominance is not an issue and the genotype can usually be deduced directly from the phenotype. However, central to classical population genetics are infinite population size, random mating, and free recombination. Consequently, as expressed by Maynard-Smith, "the alleles present at one locus are independent of those at other loci. Changes in the frequency of an allele at one locus, therefore, are independent of what is happening elsewhere in the genome: each locus can be treated individually" (Maynard-Smith, 1995). It is true that the size of bacterial populations can be practically infinite but recombination occurs extremely rarely so that changes affecting one locus can lead to the modification of others. In the succinct words of Maynard-Smith, "the genome should be treated as an inter-related whole, and not as a set of independently changing genes". The crux of the problem is knowing the exact level of recombination in bacterial populations, since "it is considerably more challenging to elaborate a theory for a population with little recombination than for one with no recombination, or a lot" (Maynard-Smith, 1995). In bacterial population genetics, sometimes we detect a degree of recombination that is too high for a pure phylogenetic approach, but too low for assessing a random interchange.

Stronger evidence for restricted recombination comes from measurements of linkage disequilibrium: that is, the tendency for particular alleles at different loci to co-occur

(Maynard-Smith et al., 1993; Haubold et al., 1998). Linkage disequilibrium (and the too-frequent occurrence of a particular combination of alleles, which is a manifestation of such disequilibrium) shows that recombination is restricted, but not absent. The determination of the relative importance of mutation in comparison with recombination is central to bacterial population genetics (Feil et al., 1999). Previous studies have demonstrated a wide variety of situations among bacterial species ranging from the clonal diversification of *Salmonella* (Selander et al., 1990) or *Escherichia coli* (Orskov et al., 1990), which are mainly due to mutation, to the frequent recombination found in *Neisseria gonorrhoeae* (O'Rourke & Stevens, 1993) or *Helicobacter pylori* (Salaun et al., 1998). Most of the population studies done with bacterial species suggest that recombination occurs in nature, and indeed may be highly important in generating variation, but that it is infrequent compared to mutation. Consequently, bacterial populations consist largely of independent clonal lineages.

The development of protein electrophoresis was a breakthrough for the study of bacterial population genetics. A pioneer in the field, Milkman (1973) used the methodology to study whether electrophoretic variation is selective or neutral. As described by Selander et al., Multilocus Enzyme Electrophoresis (MLEE) "has long been a standard method in eukaryotic population genetics and systematics" before it was applied "for studying the genetic diversity and structure in natural populations of bacteria. This research established basic population frameworks for the analysis of variation in serotypes and other phenotypic characters and has provided extensive data for systematics and useful marker systems for epidemiology" (Selander et al., 1986).

In 1998, Maiden et al. introduced Multilocus Sequence Typing (MLST), an extension of MLEE based on nucleotide sequencing that is able to determine higher levels of discrimination (more alleles per locus). In MLST "alleles are identified directly from the nucleotide sequences of internal fragments of genes rather than by comparing the electrophoretic mobilities of the enzymes they encode" (Maiden et al., 1998). In addition, this method is fully portable between laboratories and data can be stored in a single multilocus sequence database accessible via the internet (http://www.mlst.net). This approach has given "a new dimension to the elucidation of genomic relatedness at the inter- and intraspecific level by sequence analysis of housekeeping genes subject to stabilising selection. This technique has been mainly used in epidemiology, but it offers the opportunity to incorporate the insights available from population genetics and phylogenetic approaches into bacterial systematics" (Stackebrandt et al., 2002; Pérez-Losada et al., 2005; Robinson et al., 2010).

The results obtained using MLST indicate that despite the high diversity observed in bacteria, it is possible to recognize clusters with a lower degree of variation. The number of sequence types (STs) obtained is less than expected if we consider the product of the individual allelic frequencies. Nevertheless, most of the strains belong to one or a few allelic profiles, whereas most of the STs are represented by one or few strains. In addition, most STs cluster in clonal complexes constituted by closely related genotypes. Typically, each clonal complex consists of a predominant ST and a varied group of less common STs, which have different alleles only in one or two loci (Feil et al., 2004). These differences could correspond to initial stages of clonal divergence from an ancestral genotype that was the origin of the clone (Vogel et al., 2010; Willems, 2010).

2. The genus *Aeromonas*

The genus *Aeromonas* Stanier 1943 belongs to the family *Aeromonadaceae* within the class *Gammaproteobacteria* (Martin-Carnahan & Joseph, 2005). Aeromonads are autochthonous inhabitants of aquatic environments including chlorinated and polluted waters, although they can also be isolated from a wide variety of environmental and clinical sources. They cause infections in vertebrates and invertebrates, such as frogs, birds, various fish species and domestic animals. In recent years, some authors have considered *Aeromonas* as an emergent pathogen in humans, producing intestinal and extraintestinal diseases. Aeromonads are facultative anaerobic chemoorganotrophs capable of anaerobic nitrate respiration and dissimilatory metal reduction (Martin-Carnahan & Joseph, 2005).

The interest in the taxonomy of the genus *Aeromonas* has increased markedly in recent years (Janda & Abbott, 2010), and its classification, with 25 species currently recognized, remains challenging. Novel species are continuously being described, strains and species described so far are being rearranged, and DNA-DNA hybridization studies have observed discrepancies (Janda & Abbott, 2010). Historically, the genus *Aeromonas* has been divided into two groups: nonmotile, psychrophilic species, best represented by *A. salmonicida*, which are generally associated with fish diseases and motile mesophilic species associated with human diseases, including *A. hydrophila*, *A. veronii* and *A. caviae* (Martin-Carnahan & Joseph, 2005). More species have since been described and genealogies have to be adapted accordingly, although this is not always straightforward.

Bacterial species are formally defined as a group of strains that share several phenotypical characteristics and show values of DNA-DNA hybridization ≥ 70% and a 16S rRNA sequence similarity ≥ 97% with their close relatives (Stackebrandt et al., 2002). Indeed, there are hardly any examples in which strains with divergence in the 16S rRNA sequence ≤ 97% are defined as one species (Roselló-Mora & Amann, 2001). In *Aeromonas* 16S rRNA gene sequences are highly similar, being identical in some close related species such as *A. salmonicida*, *A. bestiarum, A. popoffii* and *A. piscicola*, which hampers their utility in defining species in this genus (Figueras et al., 2000; Beaz-Hidalgo et al., 2009).

Several attempts have been made to generate phylogenies using DNA gene sequences to reconstruct the correct genealogical ties among species in *Aeromonas* (Küpfer et al., 2006; Saavedra et al., 2006; Miñana-Galbis et al., 2009), but the genes chosen for this purpose are not always suitable and do not give congruent phylogenies (Farfán et al., 2010; Silver et al., 2011). Recently, two papers presenting MLST schemes for *Aeromonas* have been published (Martínez-Murcia et al., 2011; Martino et al., 2011) and there is an online MLST database of the genus *Aeromonas,* managed by Keith Jolley and curated by Barbara Cardazzo (http://pubmlst.org/aeromonas). All this accumulated data should help to establish a reliable clustering of the *Aeromonas* species and elucidate their exact boundaries.

Finally, the availability of complete genomes of different species is also useful in this task, but unfortunately in the case of *Aeromonas* only three genomes have been completed, one corresponding to the type strain of *A. hydrophila* subsp. *hydrophila* ATCC 7966 isolated from a tin of milk (Seshadri et al., 2006), the second to a fish pathogen, the strain A449 of *A. salmonicida* subsp. *salmonicida* (Reith et al., 2008) and the third to *A. veronii* isolated from an aquaculture pond sediment (Li et al., 2011). The information given by the genomes of *A. hydrophila* and *A. salmonicida* indicate that while they are of identical size (4,7Mb) and share

multiple housekeeping and virulence genes, *A. salmonicida* has acquired several mobile genetic elements, and undergone genome rearrangements and loss of genes in the process of adapting to a specific host. The genome of *A. veronii* is smaller (4,3Mb) and contains fewer virulence genes than the others.

3. "*Aeromonas hydrophila* species complex"

An example of the taxonomic complexity of the genus *Aeromonas* is the difficulty in discriminating between the phenotypically and genetically closely related species belonging to the "*Aeromonas hydrophila* species complex" (AHC), which includes: *A. hydrophila*, composed by three subspecies: A. *hydrophila* subsp. *hydrophila, A. hydrophila* subsp. *ranae* and *A. hydrophila* subsp. *dhakensis*, *A. bestiarum*, *A. popoffii* and *A. salmonicida,* divided in five subspecies: *A. salmonicida* subsp. *salmonicida*, *A. salmonicida* subsp. *masoucida*, *A. salmonicida* subsp. *achromogenes, A. salmonicida* subsp. *pectinolytica* and *A. salmonicida* subsp. *smithia* (Miñana-Galbis et al., 2002; Martin-Carnahan & Joseph, 2005). Recently, two additional species have been described in this group, *A. aquariorum* and *A. piscicola* (Martínez-Murcia et al., 2008; Beaz-Hidalgo et al., 2009). Members of the AHC were first described as strains producing the enzymes elastase, lecitinase or stapholysin (Abbott et al., 2003). They are genetically closely related and share multiple phenotypic characteristics, which makes discrimination among the species included in this group extremely difficult (Miñana-Galbis et al., 2002).

Several approaches have been used to discriminate among the AHC species: Amplified Fragment Length Polymorphisms (AFLP) (Huys et al., 1997), fluorescent AFLP (FAFLP) (Huys et al., 2002), MLEE (Miñana-Galbis et al., 2004), Random Amplified Polymorphic DNA (RAPD) and MALDI-TOF MS analysis (Martínez-Murcia et al., 2008). Although the results obtained with these methods have been useful for the taxonomy and phylogeny of the AHC, providing a hypothesis for the genealogy of strains and detailing their patterns of descent and degree of genetic variation accumulated over time, only the MLEE study has been used to elucidate their population genetic structure.

Previous studies based on the sequence analysis of several housekeeping genes have demonstrated that the AHC is not monophyletic (Soler et al., 2004; Küpfer et al., 2006; Saavedra et al., 2006; Nhung et al., 2007; Beaz-Hidalgo et al., 2009; Miñana-Galbis et al., 2009). Nevertheless, controversially, other studies have shown the monophylia of this group (Martínez-Murcia et al., 2005; Farfán et al., 2010). This conflict could be due to the incongruence of phylogenies derived from distinct gene sequence analysis.

In order to establish the population structure and divergence of the species included in this group of *Aeromonas* we studied a set of strains representative of the AHC, in which we analyzed the nucleotide sequences (total or partial) of 6 housekeeping genes: *cpn*60 (555 bp), *dna*J (891 bp), *gyr*B (1089 bp), *mdh* (936 bp), *rec*A (1065 bp), *rpo*D (843 bp), giving a total fragment length of 5379 bp.

4. Phylogenetic analysis

The relationships among the analyzed *Aeromonas* isolates are represented as a genealogical tree (Fig. 1). The tree reveals that the AHC splits into three main clusters separated by a

mean genetic distance of 0.071 (Table 1). The *A. salmonicida* cluster, although comprising five subspecies, constitutes a homogeneous clade with the lowest mean genetic distance of 0.016 (Table 1). The closest relative to *A. salmonicida*, *A. bestiarum,* has a different population structure. Rather than being a single group, it is divided in three clades, one constituted by the great majority of the *A. bestiarum* strains, a second corresponding to *A. popoffii* and the third including some *A. bestiarum* strains together with isolates of *A. piscicola.* The mean genetic distance of the *A. bestiarum* cluster is higher than in *A. salmonicida* 0.029 (Table 1). The *A. hydrophila* group, which exhibits the highest genetic distance 0.037 (Table 1), clearly separates in two clades, one including the *A. hydrophila* subsp. *hydrophila* and *A. hydrophila* subsp. *ranae* and the second constituted by *A. hydrophila* subsp. *dhakensis* and *A. aquariorum.* The presence of different clades in some of these species poses questions about the clade-species relationships.

	TN93	**Standard error**
All sequences	0.0717	0.0024
A. bestiarum	0.0286	0.0014
A. bestiarum (clade 1)	0.0157	0.0009
A. bestiarum (clade 2)	0.0084	0.0008
A. bestiarum (clade 3)	0.0100	0.0008
A. hydrophila	0.0370	0.0016
A. hydrophila (clade 1)	0.0219	0.0010
A. hydrophila (clade 2)	0.0158	0.0012
A. salmonicida	0.0163	0.0009

Table 1. Mean genetic distances for all sequences and clusters calculated using the Tamura Nei (1993) distance. Standard error estimates were obtained by the bootstrap method with 500 replicates.

	TN93	**Standard error**
A. bestiarum vs. *A. salmonicida*	0.0655	0.0033
A. bestiarum vs. *A. hydrophila*	0.1042	0.0043
A. hydrophila vs. *A. salmonicida*	0.1101	0.0049
A. bestiarum (clade 1) vs. *A. bestiarum* (clade 2)	0.0507	0.0027
A. bestiarum (clade 1) vs. *A. bestiarum* (clade 3)	0.0279	0.0021
A. bestiarum (clade 2) vs. *A. bestiarum* (clade 3)	0.0526	0.0031
A. hydrophila (clade 1) vs. *A. hydrophila* (clade 2)	0.0558	0.0030

Table 2. Mean genetic distances among different clusters obtained using the Tamura Nei (1993) distance. Standard error estimates were obtained by the bootstrap method with 500 replicates.

Estimation of distance frequency within and between clusters from multilocus sequence data provides interesting insights into the population structure of these groups. The frequency distribution of the pairwise genetic distances clearly identifies the AHC species phylogenetic structure (Fig. 2). Within the strains of the *A. bestiarum* group (Fig. 2A), the plot shows three peaks with distance values ranging from 0 to 0.063. The lowest values correspond to pairwise comparisons among isolates within clades (1, 2 and 3) and the highest to those between clades. Distances between groups allowed a clear separation of *A.*

bestiarum from the other species groups. When we plotted the pairwise distance distribution within the *A. hydrophila* group (Fig. 2B) two peaks were shown and again the lowest values are those within the clades and the highest between the clades (0-0.062). *A. hydrophila* is clearly separated from *A. salmonicida* and *A. bestiarum* despite the overlap in the graph (Fig. 2B), which is a consequence of the similar distance values between *A. hydrophila* and *A. salmonicida* (0.110) and *A. hydrophila* and *A. bestiarum* (0.104). Otherwise, species boundaries are objectively defined in the phylogenetic tree (Fig. 1). In the *A. salmonicida* group, the within distance distribution (0-0.030) appears as a single peak, as does the interspecies distance.

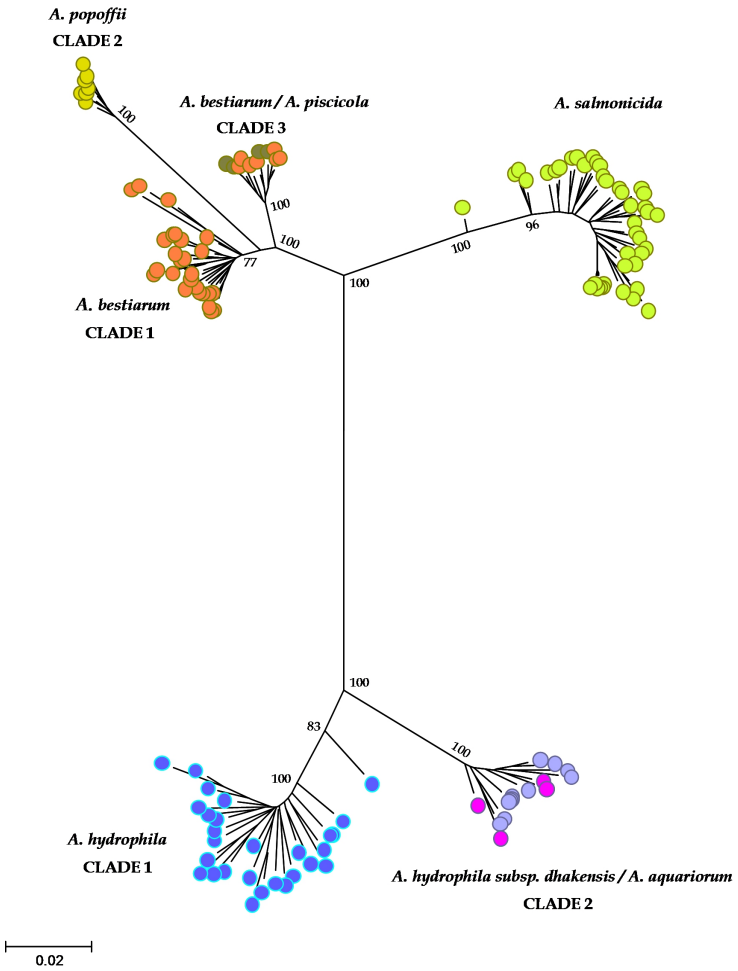

Fig. 1. Maximum likelihood tree inferred from the concatenated nucleotide sequences generated with MEGA5 software (Tamura et al., 2011). The tree was constructed using the Tamura Nei distance (TN93) and the rate of variation among sites was modelled with a gamma distribution (shape parameter = 0.5723) assuming heterogeneity of invariant sites

(best model for the data). Numbers at the branch nodes represent bootstrap values (500 replicates). The scale bar indicates the number of nucleotide substitutions per site. The GenBank/EMBL/DDBJ accession numbers of the nucleotide sequences used in this study are EU306796-EU306797, EU306804-EU306806, EU306814-EU306820, EU306822-EU306824, EU306826-EU306829, EU741625, EU741635-EU741636, EU741642, FJ936120, FJ936135, GU062399, JN711508-JN711610 (*cpn60* gene); FJ936122, FJ936136, JN215529-JN215534, JN711611-JN711731 (*dnaJ* gene); FJ936137, JN215535, JN711732-JN711858 (*gyrB* gene); HM163293-HM163294, HM163305-HM163307, HM163312-HM163318, JN660159-JN660273 (*mdh* gene); JN660274-JN660400 (*recA* gene) and EF465509-EF465510, FJ936132, FJ936138, JN215536-JN215541, JN712315-JN712433 (*rpoD* gene).

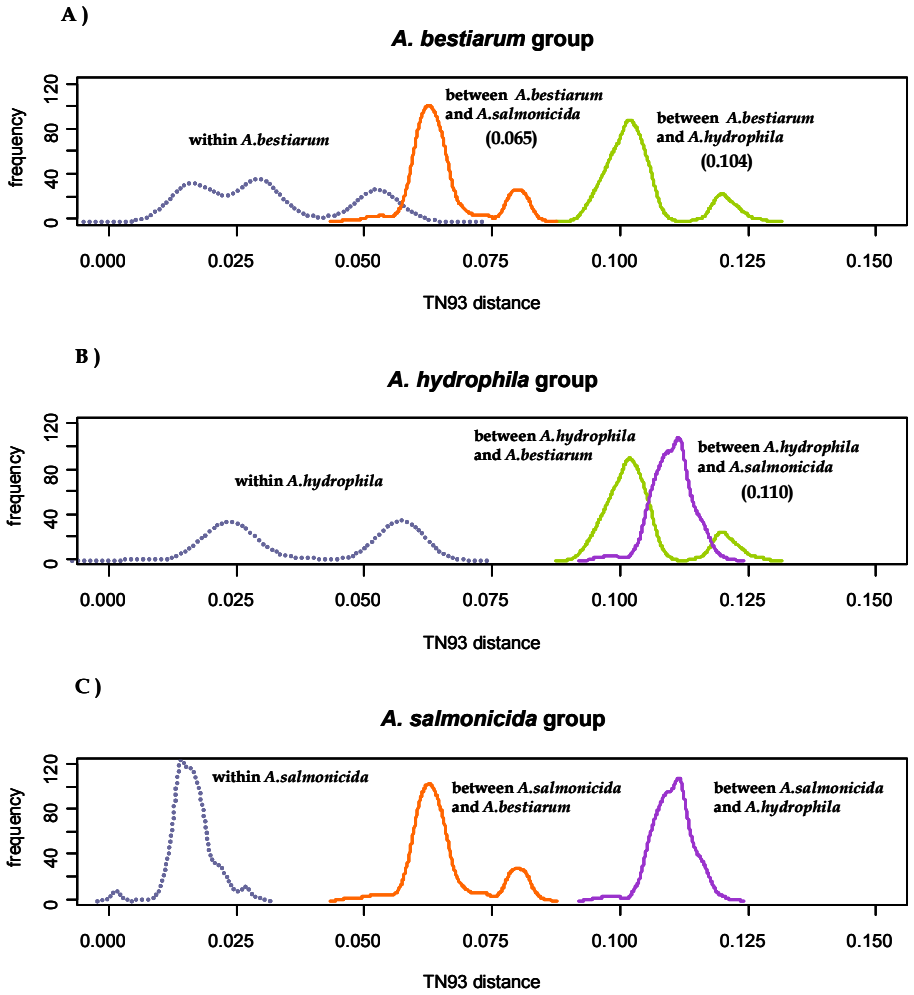

Fig. 2. Distribution of TN93 distances within and between clusters determined by pairwise comparisons. Mean distances are indicated within brackets.

Although to our knowledge there is no function that could be used to define a level of divergence for distinguishing species, it seems clear that, in bacteria, there is always an exponential relationship between interspecies recombination and sequence divergence (Roberts & Cohan, 1993; Zawadzki et al., 1995; Vulic et al., 1997). Values of divergence among clades (Table 2) seem to correspond to those of a genetically isolated biological species. The lowest value corresponds to *A. bestiarum* clade 3, indicating that its isolates are undergoing speciation, but species boundaries are not as well-defined as in the other clades.

5. Linkage disequilibrium and population structure

Linkage equilibrium is characterized by the statistical independence of alleles at all loci. A common method used in bacterial population genetics to quantify the degree of linkage between a set of loci in MLSA data is to use the index of association (I_A) (Brown et al., 1980; Maynard-Smith et al., 1993). This index compares the ratio of variance calculated from the pairwise distribution of allele mismatches in the data (V_O) and the ratio expected under a null hypothesis of linkage equilibrium (V_E). In this case

$$E(V_O) = \sum h_j (1-h_j),$$

where h_j is an unbiased estimator of the population genetic diversity equivalent to heterozygosis in diploid organisms (Table 4). The index of association (I_A) originally used by Brown et al. (1980) is computed as

$$I_A = V_O / V_E - 1,$$

which would give a value of zero if there is no association between loci (Maynard-Smith et al., 1993). Values of this index significantly different from zero reflect strong linkage disequilibrium (lower rates of recombination). The value of I_A depends on the number of loci analyzed. Haubold & Hudson (2000) subsequently proposed a standardized index of association (I_A^S) defined as

$$I_A^S = (1/l-1)\, I_A,$$

where l is the number of loci studied. This index has the advantage of being comparable between studies as long as it can be assumed that the neutral mutation parameter $\theta = 2N_e\mu$ is constant (Hudson, 1994). Two methods are commonly used for determining whether the computed I_A represents a significant deviation from linkage equilibrium ($I_A \sim 0$): resampling from randomized data sets using Monte Carlo simulations or using the parametric method described by Haubold et al. (1998). Both methods are perfectly good alternatives.

The genetic structure of the AHC, which has been previously determined using enzyme electrophoresis (MLEE), has revealed a clear clonal structure with strong linkage disequilibrium among 15 different protein loci (Miñana-Galbis et al., 2004). Additionally, this study demonstrated the usefulness of MLEE for separating the strains belonging to *A. bestiarum* and *A. salmonicida*, which are almost indistinguishable by phenotypic characteristics and 16S rRNA sequence analysis and present borderline DNA-DNA homology values.

In a multilocus sequence analysis (using six housekeeping genes) of a new set of AHC strains in which we included representatives of *A. piscicola* and *A. aquariorum*, we obtained

I_A^S values different from 0 in all cases, indicating the absence of recombination and again revealing strong linkage disequilibrium when considering both the total population and the different groups of species (Table 5). This is in spite of the high number of alleles per locus and polymorphic sites (Table 3) and huge genetic diversity (Table 4). Values of V_O, V_E, I_A^S and the 5% critical values as determined by the Monte Carlo process (L_{MC}) and from the parametric approach (L_{para}) of Haubold et al. (1998) are shown in Table 5.

Gene	Number of alleles	Number of polymorphic sites	π ± s.e.**
cpn60	95	144	0.075 ± 0.009
dnaJ	102	237	0.082 ± 0.007
gyrB	108	266	0.063 ± 0.005
*mdh**	113	225	0.059 ± 0.009
*recA**	106	282	0.071 ± 0.009
rpoD	110	212	0.090 ± 0.010

* full-length sequence gene; ** nucleotide diversity (π) ± standard error (s.e.)
Calculated by using MEGA5 software (Tamura et al., 2011).

Table 3. Sequence variation at six loci. Standard error estimates were obtained by the bootstrap method with 500 replicates.

	STs	h_j*						H ± s.e.**
		cpn60	*dnaJ*	*gyrB*	*mdh*	*recA*	*rpoD*	
All STs	127	0.9931	0.9959	0.9966	0.9978	0.9945	0.9973	0.9959 ± 0.0007
A. bestiarum	44	0.9704	0.9863	0.9863	0.9905	0.9852	0.9873	0.9843 ± 0.0029
A. hydrophila	42	0.9930	0.9884	0.9954	1.0000	0.9930	0.9930	0.9938 ± 0.0015
A. salmonicida	41	0.9744	0.9902	0.9878	0.9890	0.9707	0.9951	0.9846 ± 0.0039

* genetic diversity at individual loci (h_j); ** mean genetic diversity (H) ± standard error (s.e.)
hj = $(n/n\text{-}1)\sum p_{ij}^2$, where p_{ij} is the frequency of the ith allele at the jth locus and n the number of loci.
Data were calculated by using R statistical software (R Development Core Team, 2010).

Table 4. Genetic diversity (h) at six loci for all STs and major species sets.

	V_O	V_E	I_A^S	L_{para}	L_{MC}	P_{MC}
All STs	0.0635	0.0248	0.3131	0.0253	0.0255	< 1.00 x 10^{-04}
A. bestiarum	0.2314	0.0924	0.3010	0.0986	0.1001	< 1.00 x 10^{-04}
A. hydrophila	0.1382	0.0369	0.5495	0.0395	0.0405	< 1.00 x 10^{-04}
A. salmonicida	0.1916	0.0908	0.2222	0.0978	0.0988	< 1.00 x 10^{-04}

For the meaning of acronyms, see text. The L_{MC} and P_{MC} results were obtained from 10000 resamplings. Data were calculated by using R statistical software except for L_{para} , which was determined with the LIAN 3.5 program (Haubold & Hudson, 2000).

Table 5. Multilocus linkage disequilibrium analysis of the AHC.

During the last years, with the availability of the first DNA sequence data of individual genes, evidence of recombination at the molecular level has accumulated for *Aeromonas* in genes such as *dnaJ*, *gyrB* and *recA* (Silver et al., 2011). Incongruence between trees reconstructed from individual genes appeared as further proof of recombination at the gene level. Nevertheless, multilocus analysis with gene sequences also revealed a clear clonal

structure in this bacterial group (Table 5). The question is that although bacteria are capable of accumulating gene fragments from other bacterial species or mutations, the recombinant segments are not long enough to break the clonal structure of the population. While the absence of linkage ($I_A^S \sim 0$) is difficult to explain without assuming high levels of recombination, linkage disequilibrium does not exclude the presence of significant levels of recombination (Touchon et al., 2009).

In our study we have also determined the presence of recombinant fragments in the *recA* (in four *A. bestiarum* strains) and *dnaJ* genes (five strains, 2 *A. bestiarum*, 2 *A. hydrophila* and 2 *A. salmonicida*). However, although these strains cluster separately when the corresponding tree is constructed, revealing the different origin of the gene fragments, they group together with the other strains when a concatenated tree is generated. This confirms that recombination is not sufficient to break the genetic cohesion of this group.

6. Gene flow and divergence

The existence of barriers to gene flow such as geographical separation, ecological adaptation or the accumulation of genetic differences ultimately leads to distinct lineages. These processes are usually more complicated in prokaryotes, since the boundaries of their species are sometimes distorted by gene transfer between divergent organisms. In addition, the mechanisms that can contribute to the cohesion of groups are very different in bacteria. We have determined the divergence among the different AHC clades using the nucleotide fixation index, N_{ST} (Nei & Kumar, 2000). The use of the N_{ST} as a measure of population differentiation under certain circumstances has been criticized (Jost, 2008), but it is still used for describing the average amount of such differentiation observed from multiple locus data (Ryman & Leimar, 2009). In our study, the determination of the N_{ST} values indicated a high level of interclade genetic differentiation (N_{ST} = 0.8025). On average, most of the 80% of the total variance of nucleotide diversity was attributable to genetic differentiation among clades, whereas about 20% was found within populations. High levels of diversification were also found among most of the AHC groups of species (Table 6), with values always higher than 0.7 except in the case of *A. bestiarum* clade 3 / *A. bestiarum* clade 1 (N_{ST} = 0.5) and *A. hydrophila* clade 2 / *A. hydrophila* clade 1 (N_{ST} = 0.6). The clear divergence between the different clades described suggests they form coherent groups in which the phenomenon of recombination, if present, fails to break this consistency.

In bacterial populations it seems reasonable to equate the effect of lateral gene transfer (LTG) from other species with the product *Nm* (effective population size x migration rate) determined from N_{ST}. Indeed, assuming the limitations of the Wright island model (Wright, 1940), the value of N_{ST} at the equilibrium ($1/(2Nm+1)$) allows *Nm* to be calculated. The values obtained for all AHC clades (N_{ST} = 0.8, *Nm* ~ 0.13) again suggest that gene flow (in this case, the lateral gene transfer) is insufficient to counteract their genetic differentiation.

The McDonald-Kreitman Test (MKT, McDonald & Kreitman, 1991) was applied to our data sets to detect signs of selection (Table 7). Under neutrality, the ratio of nonsynonymous-to-synonymous fixed substitutions (between species) should be the same as the ratio of nonsynonymous-to-synonymous polymorphism (within species). We observed a high level of fixed replacements between most species studied. In all comparisons the ratio of nonsynonymous-to-synonymous substitutions was higher for fixed differences than for

polymorphisms. This result agrees with the presence of mutations under positive selection spreading quickly through a population. These changes do not contribute to polymorphism but have a cumulative effect on divergence and the fixed NS/S is consequently greater than polymorphic NS/S (Egea et al., 2008). Two of the McDonald-Kreitman tests were not significant, *A. bestiarum* versus *A. salmonicida* and *A. piscicola* (*A. bestiarum* clade 3) versus the other strains of the *A. bestiarum* group.

	***A. best* clade 1**	***A. best* clade 2**	***A. best* clade 3**	***A. hyd* clade 1**	***A. hyd* clade 2**	***A. salm***
A. bestiarum clade 1	---	0.1703	0.4532	0.1366	0.1068	0.1907
A. bestiarum clade 2	0.7460	---	0.1174	0.0863	0.0683	0.1051
A. bestiarum clade 3	0.5245	0.8099	---	0.1077	0.0833	0.1480
A. hydrophila clade 1	0.7854	0.8529	0.8228	---	0.1264	0.1264
A. hydrophila clade 2	0.8240	0.8798	0.8571	0.6428	---	0.0993
A. salmonicida	0.7239	0.8264	0.7716	0.7982	0.8343	---

Calculated using DnaSP v5.10 software (Librado & Rozas, 2009).

Table 6. Pairwise estimates of population differentiation, N_{ST} (lower-left) and gene flow, *Nm* (upper-right).

	Fixed		***Polymorphic***		***NI***	***P***
	S	**NS**	**S**	**NS**		
A. bestiarum vs. *A. hydrophila*	83	26	1235	88	0.227	0.000000 ***
A. bestiarum vs. *A. salmonicida*	57	7	964	77	0.650	0.325391 ns
A. hydrophila vs. *A. salmonicida*	118	36	1141	86	0.247	0.000000 ***
A. bestiarum (clades 1 and 3) vs. *A. bestiarum* (clade 2)	88	15	596	43	0.423	0.015757 *
A. bestiarum (clade 1 and 2) vs. *A. bestiarum* (clade 3)	21	3	659	48	0.510	0.402179 ns
A. hydrophila (clade 1) vs. *A. hydrophila* (clade 2)	56	16	803	48	0.209	0.000009 ***

Acronyms are for synonymous substitutions (S); nonsynonymous substitutions (NS); neutrality index (*NI*); *P*-value from Fisher's exact test (*P*). * $0.01<P<0.05$; ** $0.001<P<0.01$; *** $P<0.001$; not significant (ns). Calculated using DnaSP v5.10 software (Librado & Rozas, 2009).

Table 7. McDonald-Kreitman Test for molecular evidence of selection.

7. Conclusions

Developments in gene sequence analysis have greatly enhanced the study of bacterial population genetics. Gene-wide approaches to mapping bacterial diversity, which have already proved effective for gaining fresh insight into bacterial evolution, have the potential to reveal the phenotypic basis of genetic diversity in the AHC and to investigate the dynamics of this complex bacterial community. The general objective of the work described in this chapter has been to evaluate the suitability of combining population genetics and phylogenetic approaches for the delineation of bacterial species in the AHC, considered by many specialists a taxonomically tangled group.

The results obtained from the linkage disequilibrium analysis and sequence divergence show that the AHC is composed of four robust groups that basically correspond with the

phenotypically described species *A. hydrophila*, *A. bestiarum*, *A. popofii* and *A. salmonicida*. The average divergence between these clusters seems to exclude a significant influence of recombination in the genetic structure of this bacterial group and therefore they are valid taxonomic units, despite the extensive variability within some of them. Phenotypic characteristics lead to the differentiation of five *A. salmonicida* subspecies, but the lack of a consistent signal in our multilocus sequence analysis only allowed the possible differentiation of *A. salmonicida* subsp. *pectinolytica*. Similarly, it is impossible to differentiate between *A. hydrophila* subsp. *hydrophila* and *A. hydrophila* subsp. *ranae*. These results are in agreement with those obtained in a previous study with isolates belonging to the *A. veronii* Group, in which the authors failed to achieve the differentiation of biovars within these *Aeromonas* species (Silver et al., 2011). Nevertheless, *A. hydrophila* subsp. *dhakensis* strains, which cluster together with *A. aquariorum* isolates, exhibited the divergence levels of a biological species and hence deserve full species status. Consequently, the *A. aquariorum* isolates should be reclassified. Finally, in the *A. bestiarum* group we distinguished a clade (clade 3) that includes *A. piscicola* isolates as well as several strains probably misclassified as *A. bestiarum.* This clade seems to constitute an incipient new species with low values of differentiation and species boundaries less well defined than in *A. bestiarum* (clade 1) or *A. popofii* (clade 2).

It has been frequently postulated that in bacterial populations, lateral gene transfer is so common that it precludes the existence of true biological species. One of the aims of this study has been to verify if this hypothesis is applicable to our AHC data. Three lines of evidence suggest the contrary. First of all, using the Tamura Nei model, which best fits our data, we found considerable interspecific nucleotide diversity, suggesting a high degree of divergence that hampers recombination among AHC species. Secondly, the linkage disequilibrium analysis of six loci reveals a strong disequilibrium with I_A^S values, suggesting little or null influence of recombination in the genetic structure of AHC species. Thirdly, the N_{ST} values obtained reflect a high degree of differentiation between clades. In short, the genetic structure of the AHC appears to confirm that the entities phenotypically described as species form cohesive groups in which genetic recombination plays a limited role in reducing genetic variation and can be defined as biological species.

Like other authors (Lan & Reeves, 2001; Vinuesa et al., 2005), we agree that a combination of phylogenetic and population genetic studies is currently the best theoretical and practical approach to delineate species as natural and discrete lineages in the bacterial world.

8. Acknowledgments

We are very grateful to Katri Berg (University of Helsinki, Finland), Margarita Gomila (Universitat de les Illes Balears, Mallorca, Spain) and Carmen Gallegos (Hospital Sant Llàtzer, Mallorca, Spain) for kindly providing some of the strains used in this study. This research was supported by the project CGL-2008-03281/BOS from the Ministerio de Educación y Ciencia, Spain.

9. References

Abbott, L.S.; Cheung, W.K.W. & Janda, M.J. (2003). The genus *Aeromonas*: Biochemical characteristics, atypical reactions, and phenotypic identification schemes. *Journal of Clinical Microbiology,* Vol. 41, No. 6, pp. 2348-2357.

Beaz-Hidalgo, R.; Alperi, A.; Figueras, M.J. & Romalde, J.L. (2009). *Aeromonas piscicola* sp. nov., isolated from diseased fish. *Systematic and Applied Microbiology,* Vol. 32, No. 7, pp. 471-479.

Brown, A.D.; Feldman, M.W. & Nevo, E. (1980). Multilocus structure of natural populations of *Hordeum spontaneum*. *Genetics,* Vol. 96, No. 2, pp. 523-536.

Egea, R.; Casillas, S. & Barbadilla, A. (2008). Standard and generalized McDonald-Kreitman test: a website to detect selection by comparing different classes of DNA sites. *Nucleic Acids Research,* Vol. 36, No. 2, pp. W157-W162.

Farfán, M.; Miñana-Galbis, D.; Garreta, A.; Lorén, J.G. & Fusté, M.C. (2010). Malate dehydrogenase: a useful phylogenetic marker for the genus *Aeromonas*. *Systematic and Applied Microbiology,* Vol. 33, No. 8, pp. 427-435.

Feil, E.J.; Maiden, M.C.J.; Achtman, M. & Spratt, B.G. (1999). The relative contributions of recombination and mutation to the divergence of clones of *Neisseria meningitidis*. *Molecular Biology and Evolution,* Vol. 16, No. 11, pp. 1496-1502.

Feil, E.J.; LI, B.C.; Aanensen, D.M.; Hanage, W.P. & Spratt, B.G. (2004). eBURST: Inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus typing data. *Journal of Bacteriology,* Vol. 186, No. 5, pp. 1518-1530.

Figueras, M.J.; Soler, L.; Chacón, M.R.; Guarro, J. & Martínez-Murcia, A.J. (2000). Extended method for discrimination of *Aeromonas* spp. by 16S rDNA RFLP analysis. *International Journal of Systematic and Evolutionary Microbiology,* Vol. 50, No. 6, pp. 2069-2073.

Janda, J.M. & Abbott, S. (2010). The genus *Aeromonas:* taxonomy, pathogenicity, and infection. *Clinical Microbiology Reviews,* Vol. 23, No. 1, pp. 35-73.

Jost, L. (2008). G_{ST} and its relatives do not measure differentiation. *Molecular Evolution*, Vol. 17, No. 18, pp. 4015-4026.

Haubold, B.; Travisano, M.; Rainey, P.B. & Hudson, R.R. (1998). Detecting linkage disequilibrium in bacterial populations. *Genetics,* Vol. 150, No. 4, pp. 1341-1348.

Haubold, B. & Hudson, R.R. (2000). LIAN 3.0: detecting linkage disequilibrium in multilocus data. *Bioinformatics,* Vol. 16, No. 9, pp. 847-848.

Hudson, R.R. (1994). Analytical results concerning linkage disequilibrium in models with genetic transformation and conjugation. *Journal of Evolutionary Biology,* Vol. 7, No. 5, pp. 535-548.

Huys, G.; Kämpfer, P.; Altwegg, M.; Kersters, I.; Lamb, A.; Coopman, R.; Lüthy-Hottenstein, J.; Vancanneyt, M.; Janssen, P. & Kersters, K. (1997). *Aeromonas popoffii* sp. nov., a mesophilic bacterium isolated from drinking water production plants and reservoirs. *International Journal of Systematic Bacteriology,* Vol. 47, No. 4, pp. 1165-1171.

Huys, G.; Kämpfer, P.; Albert, M.J.; Kühn, I.; Denys, R. & Swings, J. (2002). *Aeromonas hydrophila* subsp. *dhakensis* subsp. nov., isolated from children with diarrhoea in Bangladesh, and extended description of *Aeromonas hydrophila* subsp. *hydrophila* (Chester 1901) Stanier 1943 (Approved Lists 1980). *International Journal of Systematic Bacteriology,* Vol. 52, No. 3, pp. 705-712.

Küpfer, M.; Kuhnert, P.; Korczak, B.M.; Peduzzi, R. & Demarta, A. (2006). Genetic relationships of *Aeromonas* strains inferred from 16S rRNA, *gyrB* and *rpoB* gene sequences. *International Journal of Systematic and Evolutionary Microbiology,* Vol. 56, No. 12, pp. 2743-2751.

Lan, R. & Reeves, P.R. (2001). When does a clone deserve a name? A perspective on bacterial species based on population genetics. *Trends in Microbiology,* Vol. 9, No. 9, pp. 419-424.

Levin, B.R. & Bergstrom, C.T. (2000). Bacteria are different: observations, interpretations, speculations, and opinions about the mechanisms of adaptive evolution in

prokaryotes. *Proceedings of the National Academy of Sciences of the United States of America,* Vol. 97, No. 13, pp. 6981-6985.

Li, Y.; Liu, Y.; Zhou, Z.; Huang, H.; Ren, Y.; Zhang, Y.; Li, G.; Zhou, Z. & Wang, L. (2011). Complete genome sequence of *Aeromonas veronii* strain B565. *Journal of Bacteriology,* Vol. 193, No. 13, pp. 3389-3390.

Librado, P. & Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics,* Vol. 25, No. 11, pp. 1451-1452.

Maiden, M.C.; Bygraves, J.A.; Feil, E.; Morelli, G.; Russell, J.E.; Urwin, R.; Zhang, Q.; Zhou, J.; Zurth, K.; Caugant, D.A.; Feavers, I.M.; Achtman, M. & Spratt, B.G. (1998). Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proceedings of the National Academy of Sciences of the United States of America,* Vol. 95, No. 6, pp. 3140-3145.

Martin-Carnahan, A. & Joseph, S.W. (2005). Genus I. *Aeromonas* Stanier 1943, 213^{AL}, In: *Bergey's Manual of Systematic Bacteriology,* Garrity, G.M.; Brenner, D.J.; Krieg, N.R. & Staley, J.T. (eds.), Vol. 2, part B, pp. 557-578, Springer, New York.

Martínez-Murcia, A.J.; Soler, L.; Saavedra, M.J.; Chacón, M.R.; Guarro, J.; Stackebrandt, E. & Figueras, M.J. (2005). Phenotypic, genotypic, and phylogenetic discrepancies to differentiate *Aeromonas salmonicida* from *Aeromonas bestiarum*. *International Microbiology,* Vol. 8, No. 4, pp. 259-269.

Martínez-Murcia, A.J.; Saavedra, M.J.; Mota, V.R.; Maier, T.; Stackebrandt, E. & Cousin, S. (2008). *Aeromonas aquariorum* sp. nov., isolated from aquaria of ornamental fish. *International Journal of Systematic and Evolutionary Microbiology,* Vol. 58, No. 5, pp. 1169-1175.

Martínez-Murcia, A.J.; Monera, A.; Saavedra, M.J.; Oncina, R.; López-Alvarez, M.; Lara, E. & Figueras, M.J. (2011). Multilocus phylogenetic analysis of the genus *Aeromonas*. *Systematic and Applied Microbiology,* Vol. 34, No. 3, pp. 189-199.

Martino, M.E.; Fasolato, L.; Montemurro, F.; Rosteghin, M.; Manfrin, A.; Patarnello, T.; Novelli, E. & Cardazzo, B. (2011). Determination of microbial diversity of *Aeromonas* strains on the basis of multilocus sequence typing, phenotype, and presence of putative virulence genes. *Applied and Environmental Microbiology,* Vol. 77, No. 14, pp. 4986-5000.

Maynard-Smith, J. (1991). The population genetics of bacteria. *Proceedings of the Royal Society B Biological Sciences,* Vol. 245, No. 1312, pp. 37-41.

Maynard-Smith, J.; Smith, N.H.; O'Rourke, M. & Spratt, B.G. (1993). How clonal are bacteria? *Proceedings of the National Academy of Sciences of the United States of America,* Vol. 90, No. 10, pp. 4384-4388.

Maynard-Smith, J. (1995). Do bacteria have population genetics?, In: *Population Genetics of Bacteria,* Baumberg, S.; Young, J.P.W; Wellington, E.M.H. & Saunders, J.R. (eds.), pp. 1-12, Cambridge University Press, Cambridge.

McDonald, J.H. & Kreitman, M. (1991). Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature,* Vol. 351, pp. 652-654.

Milkman, R. (1973). Electrophoretic variation in *Escherichia coli* from natural sources. *Science,* Vol. 182, No. 116, pp. 1024-1026.

Miñana-Galbis, D.; Farfán, M.; Lorén, J.G. & Fusté, M.C. (2002). Biochemical identification and numerical taxonomy of *Aeromonas* spp. isolated from environmental and clinical samples in Spain. *Journal of Applied Microbiology,* Vol. 93, No. 3, pp. 420-430.

Miñana-Galbis, D.; Farfán, M.; Fusté, M.C. & Lorén, J.G. (2004). Genetic diversity and population structure of *Aeromonas hydrophila*, *Aer. bestiarum*, *Aer. salmonicida* and

Aer. popoffii by multilocus enzyme electrophoresis (MLEE). *Environmental Microbiology,* Vol. 6, No. 3, pp. 198-208.

Miñana-Galbis, D.; Urbizu-Serrano, A.; Farfán, M.; Fusté, M.C. & Lorén, J.G. (2009). Phylogenetic analysis and identification of *Aeromonas* species based on sequencing of the *cpn60* universal target. *International Journal of Systematic and Evolutionary Microbiology,* Vol. 59, No. 8, pp. 1976-1983.

Nei, M., & Kumar, S. (2000). Genetic polymorphism and evolution, In: *Molecular Evolution and Phylogenetics,* pp. 231-264, Oxford University Press, New York.

Nhung, P.H.; Hata, H.; Ohkusu, K.; Noda, M.; Shah, M.M.; Goto, K. & Ezaki, T. (2007). Use of the novel phylogenetic marker *dnaJ* and DNA-DNA hybridization to clarify interrelationships within the genus *Aeromonas*. *International Journal of Systematic and Evolutionary Microbiology,* Vol. 57, No. 6, pp. 1232-1237.

O'Rourke, M. & Stevens, E. (1993). Genetic structure of *Neisseria gonorrhoeae* populations: a non-clonal pathogen. *Journal of General Microbiology,* Vol. 139, No. 11, pp. 2603-2611.

Orskov, F.; Whittam, T.S.; Cravioto, A. & Orskov, I. (1990). Clonal relationships among classic enteropathogenic *Escherichia coli* (EPEC) belonging to different O groups. *The Journal of Infectious Diseases,* Vol. 162, No. 1, pp. 76-81.

Pérez-Losada, M.; Brownw, E.B.; Madsen, A.; Wirth, T.; Viscidi, R.P. & Crandall, K.A. (2006). Population genetics of microbial pathogens estimated from multilocus sequence typing (MLST) data. *Infection, Genetics and Evolution,* Vol. 6, No. 2, pp. 97-112.

R Development Core Team (2010). *R: a language and environment for statistical computing,* R Foundation for Statistical Computing, Vienna, Austria, URL http://www.r-project.org/.

Reith, M.E.; Singh, R.K.; Curtis, B.; Boyd, J.M.; Bouevitch, A.; Kimball, J.; Munholland, J.; Murphy, C.; Sarty, D.; Williams, J.; Nash, J.H.E.; Johnson, S.C. & Brown, L.L. (2008). The genome of *Aeromonas salmonicida* subsp. *salmonicida* A449: insights into the evolution of a fish pathogen. *BMC Genomics,* Vol. 9, No. 427.

Roberts, M.S. & Cohan, F.M. (1993). The effect of DNA sequence divergence on sexual isolation in *Bacillus*. *Genetics,* Vol. 134, No. 2, pp. 401-408.

Robinson, A.; Falush, D. & Feil, E.J. (eds.) (2010). *Bacterial Population Genetics in Infectious Disease*, Wiley & Sons, ISBN 978-0-470-42474-2, New Yersey.

Roselló-Mora, R. & Amann, R. (2001). The species concept for prokaryotes. *FEMS Microbiology Reviews*, Vol. 25, No. 1, pp. 39-67.

Ryman, N. & Leimar, O. (2009). G_{ST} is still a useful measure of genetic differentiation – a comment on Jost's D. *Molecular Ecology*. Vol. 18, No. 10, pp. 2084-2087.

Saavedra, M.J.; Figueras, M.J. & Martínez-Murcia, A.J. (2006). Updated phylogeny of the genus *Aeromonas*. *International Journal of Systematic and Evolutionary Microbiology,* Vol. 56, No. 10, pp. 2481-2487.

Salaun, L.; Audibert, C.; Le Lay, G.; Burocoa, C.; Fauchère, J.L. & Picard, B. (1998). Panmintic structure of *Helicobacter pylori* demonstrated by the comparative study of six genetic markers. *FEMS Microbiology Letters,* Vol. 161, No. 2, pp. 231-239.

Seshadri, R.; Joseph, S.W.; Chopra, A.K.; Sha, J.; Shaw, J.; Graf, J.; Haft, D.; Wu, M.; Ren, Q.; Rosovitz, M.J.; Madupu, R.; Tallon, L.; Kim, M.; Jin, S.; Vuong, H.; Stine, O.C.; Ali, A.; Horneman, A.J. & Heidelberg, J.F. (2006). Genome sequence of *Aeromonas hydrophila* ATCC 7966: jack of all trades. *Journal of Bacteriology,* Vol. 188, No. 23, pp. 8272-8282.

Selander, R.K.; Caugant, D.A.; Ochman, H.; Musser, J.M.; Gilmour, M.N. & Whittam, T.S. (1986). Methods of multilocus enzyme electrophoresis for bacterial population

genetics and systematics. *Applied and Environmental Microbiology,* Vol. 51, No. 5, pp. 873-884.

Selander, R.K.; Beltran, P.; Smith, N.H.; Helmuth, R.; Rubin, F.A.; Kopecko, D.J.; Ferris, K.; Tall, B.D.; Cravioto, A. & Musser, J.M. (1990). Evolutionary genetic relationships of clones of *Salmonella* serovars that cause human typhoid and other enteric fevers. *Infection* & *Immunity,* Vol. 58, No. 7, pp. 2262-2275.

Silver, A.C.; Williams, D.; Faucher, J.; Horneman, A.J.; Gogarten, J.P. & Graf, J. (2011). Complex evolutionary history of the *Aeromonas veronii* group revealed by host interaction and DNA sequence data. *PLoS ONE,* Vol. 6, No. 2, e16751.

Soler, L.; Yañez, M.A.; Chacón, M.R.; Aguilera-Arreola, M.G.; Catalán, V.; Figueras, M.J. & Martínez-Murcia, A.J. (2004). Phylogenetic analysis of the genus *Aeromonas* based on two housekeeping genes. *International Journal of Systematic and Evolutionary Microbiology,* Vol. 54, No. 5, 1511-1519.

Stackebrandt, E.; Frederiksen, W.; Garrity, G.M.; Grimont, P.A.D.; Kämpfer, P.; Maiden, M.C.J.; Nesme, X.; Rosselló-Mora, R.; Swings, J.; Trüper, H.G.; Vauterin, L.; Ward, A.C. & Whitman, W.B. (2002). Report of the *ad hoc* committee for the re-evaluation of the species definition in bacteriology. *International Journal of Systematic and Evolutionary Microbiology,* Vol. 52, No. 3, 1043-1047.

Tamura, K.; Peterson, D.; Peterson, N.; Stecher, G.; Nei, M. & Kumar, S. (2011). MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution,* DOI: 10.1093/molbev/msr121.

Touchon, M.; Hoede, C.; Tenaillon, O.; Barbe, V.; Baeriswy, S.; Bidet, P.; Bingen, E.; Bonacorsi, S.; Bouchier, C.; Bouvet, O.; Calteau, A.; Chiapello, H.; Clermont, O.; Cruveiller, S.; Danchin, A.; Diard, M.; Dossat, C.; Karoui, M.E.; Frapy, E.; Garry, L.; Ghigo, J.M.; Gilles, A.M.; Johnson, J.; LeBouguénec, C.; Lescat, M.; Mangenot, S.; Martínez-Jéhanne, V.; Matic, I.; Nassif, X.; Oztas, S.; Petit, M.A.; Pichon, C.; Rouy, Z.; Ruf, C.S.; Schneider, D.; Tourret, J.; Vacherie, B.; Vallenet, D.; Médigue, C.; Rocha, E.P.C. & Denamur, E. (2009). Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS Genetics,* Vol. 5, No. 1, e1000344.

Vinuesa, P.; Silva, C.; Werner, D. & Martínez-Romero, E. (2005). Population genetics and phylogenetic inference in bacterial molecular systematics: the roles of migration and recombination in *Bradyrhizobium* species cohesion and delineation. *Molecular Phylogenetics and Evolution,* Vol. 34, No. 1, pp. 29-54.

Vogel, U., Schoen, C. & Elias, J. (2010). Population genetics of *Neisseria meningitidis*. In: *Bacterial Population Genetics in Infectious Disease,* Robinson, A.; Falush, D. & Feil, E.J. (eds.), pp. 247-267, Wiley & Sons, ISBN 978-0-470-42474-2, New Yersey.

Vulic, M.; Dionisio, F.; Taddei, F. & Radman, M. (1997). Molecular keys to speciation: DNA polymorphism and the control of genetic exchange in enterobacteria. *Proceedings of the National Academy of Sciences of the United States of America,* Vol. 94, No. 18, pp. 9763-9767.

Willems, R.J. (2010). Population genetics of *Enterococcus*. In: *Bacterial Population Genetics in Infectious Disease,* Robinson, A.; Falush, D. & Feil, E.J. (eds.), pp. 195-216, Wiley & Sons, ISBN 978-0-470-42474-2, New Yersey.

Wright, S. (1940). Breeding structure of populations in relation to speciation. *American Naturalist,* Vol. 74, No. 752, pp. 232-248.

Zawadzki, P.; Roberts, M.S. & Cohan, F.M. (1995). The log-linear relationship between sexual isolation and sequence divergence in *Bacillus* transformation is robust. *Genetics,* Vol. 140, No. 3, pp. 917-932.