INT. J. BIOL. BIOTECH., 4 (1): 15-18, 2007.

EXPLORING VERTICAL AND HORIZONTAL GENE TRANSFER IN *PASTEURELLA MULTOCIDA* PM70 BY COMPARATIVE GENOMICS STUDIES

Mushtaq Hussain M. Tanweer Khan and Abdul Wajid

Institute of Biotechnology and Genetic Engineering (KIBGE), University of Karachi-75270, Pakistan.

ABSTRACT

The present study has been conducted to explore different patterns and modes of *Pasteurella multocida* evolution. Total protein of *P.multocida* Pm70 with genomic synteny were subjected to total best protein identity comparison to all sequenced microbes at cutoff value of identity 70%; similarity 70% and *p*-value 1. Proteome identity of test and closely matched organisms lies in the range of 23.9% to 0.72%. The findings were further analyzed by comparative 16srDNA sequence alignment. Entire nodal branch, leading to *P.multocida* was selected for further analysis. Closest relatives thus found belongs to the species of genera *Mannheimia* and *Hemophilus* which were compared for C-Level paradox, AT content, purine stretches, simple repeats, gene density, origin and terminal of replication etc. Fourier transformation was applied to transform entire DNA of test and phylogenetically related microbes into numerical values. From mentioned digitization, stacking energy, protein deformability, propeller twists were deduced; this provided conspicuous evidences of potential vertical and lateral gene transfer in *P.multocida*. Additionally, similarity in the pattern of pseudo 2D gel profile among the microbes present in targeted nodal branch distinguished between physico-chemical nature of horizontally and vertically transferred genes

Key Words: Microbial evolution, *P.multocida*, Fourier transformation, Lateral gene transfer.

INTRODUCTION

Bacterial is one of the perplexing phenomena. In spite of many notions, two hypotheses are very popular and enjoy strong support and appreciation. Earlier models of understanding microbial adaptation, evolution and speciation are primarily focused on clonality and periodic selection (Levin, 1981). However advent in whole genome sequencing, the role of horizontal gene transfer (HGT) in bacterial evolution and diversification has been profoundly delineated (Gogarten *et al*., 2002). Despite previous suggestions of ignorable rate of homologous recombination and so their function in bacterial evolution events (Cohan ,1994a, 1994b), whole genome sequencing along with arsenals of dinucleotides and Markov chain analysis not only severely challenged clonality mediated evolution dictum but also explained unorthodox behavior which occasionally appeared in phylogenetic tree based on 16srRNA sequences (Lawrence and Ochman, 1998; Feil *et al*., 2001; Coenye *et al*., 2005; Doolittle, 2005). Besides variation in natural tendency to uptake the foreign DNA (Bakkali *et al*., 2004) several other sequential features like AT content, simple repeats, global repeats (Watson and Crick repeats) and symmetry elements percentage and structural attributes for instance propeller twist, stacking energy and protein induced deformability of genome selects route for successful HGT (ElHassan and Calladine, 1996; Olson *et al*., 1998; Sinden *et al*., 1998; Jensen *et al*., 1999; Worning *et al*., 2000).

P.multocida, a common nasooropharyngeal flora of many pet animals is widely accepted as an etiological agent of fowl cholera, hemorrhagic septicemia in cattles, conjunctivitis, nystalgia, trochilitis and cancers in humans (Rhoades and Rimler, 1991; May *et al*., 2001).

In this paper we established total protein and 16srRNA based phylogeny of *P. multocida* Pm70 and analyzed the role of HGT and VGT (Vertical gene transfer) in evolution of *P. multocida* Pm70 genome.

MATERIALS AND METHODS

Phylogeny Establishment: *P.multocida* Pm70 16srRNA sequences were downloaded from NCBI genome server. Genome size, number of ORFs, and total repeats and mobile elements were retrieved from TIGR server. The genome of *P.multocida* Pm70 contains six paralogs of 16srRNA genes ranging from 341430-2920, 541959-3449, 1081881-0390, 1690577-2068, 1761354-2844 and 1942771-1281 bp. All paralogs genes were aligned by Clustal X alignment software, 16srRNA gene sequence located from 341430-2920 was selected for non-redundant blast alignment on NCBI server (Altschul *et al*., 1997). Out of 1000 selected, entries of different genera with closest homology were downloaded and aligned by Clustal X. Phylogenetic distances were calculated with phylogeny inference program Phylip tree view (Felsenstein, 1993). Total proteome identity and pseudo 2D gel of the organisms were deduced from TIGR genomic server (www.tigr.org) with a threshold value of 70% similarity and identity and *p value*=1.

Periodicity Calculation: Genomes of phylogenetically closed organisms were subjected for periodicity calculations using CBS online server. Periodicity is the function of AT content, stacking energy, propeller twist and protein deformability which was deduced by autocorrelation functions and subsequently subjected to Fourier Transformations using method described by Olson *et al*., 1998 and Worning *et al.,* 2000.

RESULTS AND DISCUSSION

Total protein identity analysis showed 23.9% identity of *Manneheimia succiniciproducens* (MS) with *Pasteurella multocida* (PM). While 20.48% and 17.87% identity were found in *Hemophilus influenza* NP (HIN) and *Hemophilus influenza* Rd (HIR) respectively to PM. Only 3.67 % of *Hemophilus ducreyi* (HD) proteome has appeared identical to PM. Rest of the sequenced microbes which mainly includes *Escherichia coli* K12-MG1655, *Yersinia Pestis biovar Medievalis, Vibrio cholerae* EITor N16961, *Pseudomonas aeruginosa* PAO1, *Neisseria meningitidis* MC58, *Legionella pneumophilia* Lens, *Brucella melitensis* 16M showed less than 1% of total protein identity with PM **(Table 1)**. Phylogenetic tree based on 16srRNA sequences alignments by neighbor joining method segregate into three major nodal branches, two of them leads to gamma proteobacteria while remaining branch further disintegrate into alpha and beta proteobacteria. Organisms having more than 1% of total protein identity to PM immediately dissociate from the rest of gamma proteobacteria. In the nodal branch leading to PM first radiated the MS and HD subsequently further cleavage appeared on remaining branch length that causes divergence of PM, HIN and HIR **(Fig.1.)**. On gross scale phylogeny based on 16srRNA sequences supports the predicted evolution based on total proteome identity, as with high degree of protein identity among MS, PM and Hemophilus species, the very immediately disintegrated from the main tree but all stays on same nodal branch on phylogenetic tree. Nodal Branch length suggests that the mentioned radiation has occurred around 680mya at Cambarian geological era (May *et al*., 2001; Battistuzzi *et al*., 2004). From early off shooting of MS on nodal branch leads to PM that makes it (MS) closer to ancestors and suggests that it contains majority of core gene setup that is mandatory for the survival of microbes having at least same ancestral origin. Later radiation in the evolutionary course occurs around 270mya (Permian era), which leads to the divergence of PM and HI strains as deduced from the branch length. Though present on the same nodal branch, the position of MS is very much unprecedented in the light of its highest protein identity of 23.9% with PM due to its profound distance. As the microbial evolution is the function of various factors which primarily includes vertical gene transfer accompanying with variable mutation rates and horizontal gene transfer, hence any of these factors could lead to this anomalous radiation of MS and or PM on phylogenetic tree in light of high protein identity (May *et al*., 2001; Feil *et al*., 2001; Coenye *et al*., 2005; Doolittle, 2005).

Table 1 showed comparative values of genome size, number of genes, GC and AT contents and various repeats percentages in genome. Except for the GC and AT skew, and coding density of the genome, PM and MS share almost similar values. Additionally genome size, number of genes, origin of replication, repeats like global direct, global inverted, local direct, local inverted, local mirror and local inverted show close resemblance between MS and PM. Such close match conspicuously indicates the transfer of ancestral genes via vertical gene transfer (Gogarten *et al*., 2002). In contrast to this discrepancies among mentioned attributes between *Hemophilus* species and PM could be inferred in terms of horizontal gene transfer (Ochman *et al*., 2000; Lawrence and Hendrickson, 2005).

Organisms	P.I in %	G.S in MBp	C.D in %	Or iR	Ter	Repeats in %							Contents in %	
						GD	GI	LD	LI	LM	LE	SP	AT	GC
M.succiniciproducens	23.9	2.31	89.3	333	1853	3.3	3.1	5.3	4.7	3.7	3.8	0.4	57.5	42.5
H.ducreyi	3.67	1.7	84.8	0	1138	4.8	3.8	59	4.1	4.6	3.1	0.6	61.8	38.2
P. multocida	100	2.31	88.9	171 1	915	3.1	1.7	4.8	4.1	3.1	3.1	0.3	59.1	40.6
H.influenza NP	20.48	1.91	86.8	774	1620	3.9	3.5	1.3	5.1	4.5	3.3	0.5	61.9	38.1
H.influenza Rd	17.87	1.83	85.5	334	1476	3.6	3.8	1.3	5.1	4.6	3.6	0.5	61.9	38.1

Table.1. Comparative account of various genetic attributes of five bacteria present on nodal branch leading to *P.multocida.* Note the similarity between different attributes between PM and MS as mentioned in text. P.I. (Protein identity), G.S. (Genome Size), C.D. (Coding density), OriR (Origin of replication), Ter (Termination of replication), GD (Global Direct), GI (Gloal Inverted), LD (Local Direct), LI (Local Inverted), LM (Local Mirror).

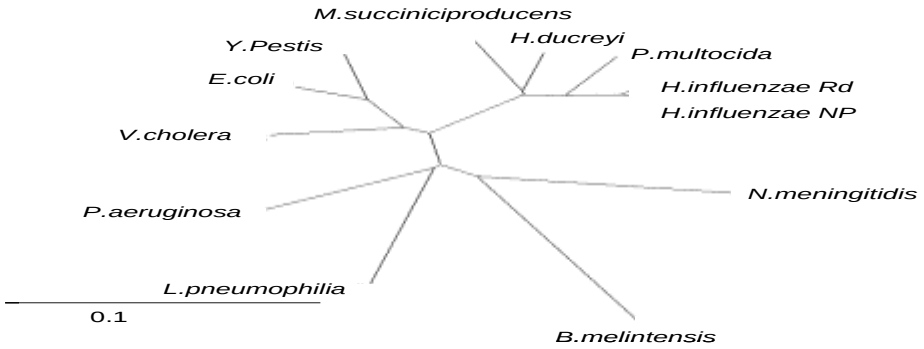

Fig.1. Phylogenetic lineage of Proteobacteria: 16srRNA gene sequence based phylogenetic tree constructed using neighbor joining distance method using ClustalX and Phylip.

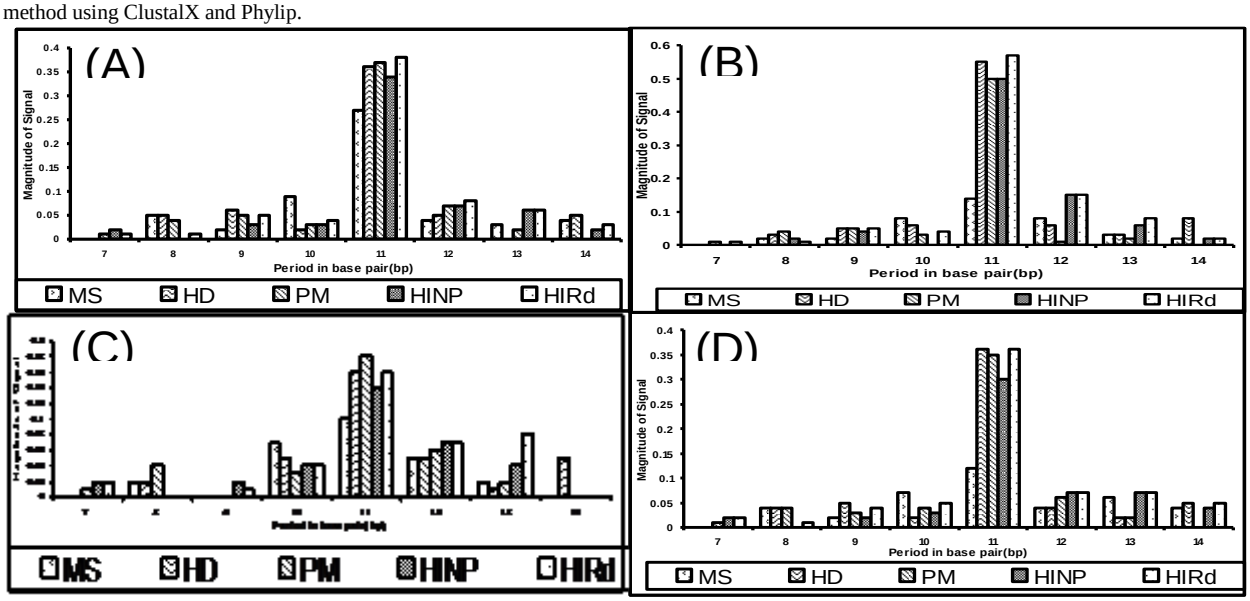

Fig.2. Comparative time series bar diagram for periodicity in bacteria: Note the strong signal amplitude in case of *P.multocida* and *H..influenzae* strains at 11bp in all cases renders to strong DNA flexibility. **KEY:** AT content (A), Propeller twist (B), Protein deformability (C), Stacking Energy (D), *M. succiniciproducens* (MS), *H.ducreyi* (HD), *P.multocida* (PM), *H.influenzae* NP (HINP) and *H..influenzae* Rd (HIRd).

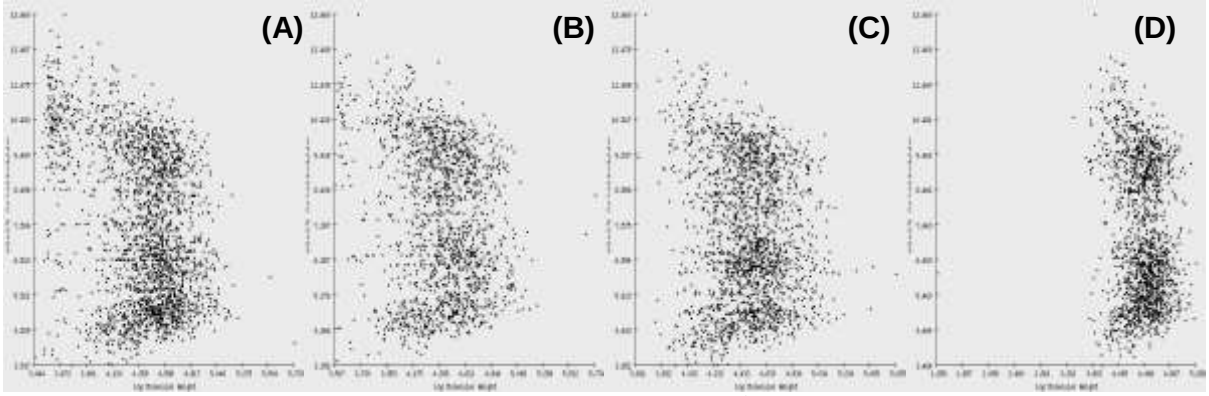

Fig.3. Pseudo 2D gel electrophoresis: Pseudo 2D gel electrophoresis banding patterns from A to D (M.succinoproducens; H.duceryi; P.multocida; H.influenzeae, respectively) clearly indicates a retrograde evolution in terms of low molecular weight neutral to alkaline proteins while progressive evolution in high molecular weight proteins of acidic to neutral proteins. The sharing band patterns (double headed arrow) between C and D reveals the nature and tentative functions of the genes transferred via HGT. On contrary broken arrow between A and C indicates core gene transfer by vertical means.

The spectra of auto correlated and subsequently Fourier transformed DNA sequences of all analyzed genomes for three different structural parameters like propeller twist, protein induced deformability and stacking energy and one sequential feature AT content showed one distinct peak well above noise level located between 10bp to 11bp regions except in case of MS and HRD where protein induced deformability showed a secondary peaks around 8bp and 12.9bp respectively **(Fig.2.)**. The spectra are

characteristically close to the previous finding of Worning *et al*., 2000 suggesting its bacterial origin. Moreover the strong signals among PM and HI indicate greater DNA flexibility, which renders the increased potential of genetic recombination; hence, conclusively described the strong possibility of horizontal gene transfer between PM and HI. Moreover relative to PM and Hemophilus, MS showed low amplitude signals, which turn into its low DNA flexibility. Such rigidity decreases the potential of organism to be involved in any mode of HGT. Taking evolutionary course appeared on the nodal branch as a time series process Pseudo 2D protein gel patterns showed gradual depletion of low molecular weight protein of relatively high isoelectric point (pI). This gradual disappearance is concomitant to the appearance of protein of acidic to neutral pI values **(Fig.3.)**. This emergence of new protein might be the reflection of climatic shifts during Cambarian and Permian geological era. More importantly it could be explained evolutionary in terms of adaptive response of organisms or its ancestors to encounter the changing environmental features. Additionally, close similarity between the banding pattern of PM and MS validates the standpoint of ancestral gene persistence between the two. However some homologous coordinates in Pseudo 2D gel between HI strains and PM particularly around the region of high molecular weight with neutral pI exhibit nature of horizontally transferred genes. Sequence of such coordinates showed strong homology with the dispensable genes like of glycerol kinase, oligopeptide ABC transporters, chemo-taxis mediating molecules and virulence factors like hemeaglutinin. The findings further encourage the chances of these genes to be exchanged laterally. evolution of *P.multocida* is the function of both ancestral gene(s) persistence via vertical gene transfer with mutability and horizontal gene transfer. Transfer of 16srRNA gene despite of its core gene nature might occur through HGT from *H.influenzae* to *P.multocida*. However, it is our belief that the present work only adumbrates the evolutionary course of *P.multocida*. Further in depth analysis with reference to individual gene/protein identity and more importantly dissimilarity using additional informatics tools is still under way and will be reported elsewhere.

REFERENCES

Altschul, S.F., T.L. Madden, A.A. Scaffer, J. Zhang, Z. Zhang, W. Miller and D.J. Lipman (1997). Basic local alignment search tools. *Nucleic Acid Res*., 25: 3389-3402.

Felsenstein, J. (1993). *PHYLIP program*, Dept of Genetics, Univ. of Washington.

Bakkali, M., T.Y.Chen, H.C. Lee, and R. J. Redfield (2004). Evolutionary stability of DNA uptake signal sequences in Pateurellaceae. *Proc. Natl. Acad. Sci. USA*., 101: 4513-4518.

Battistuzzi, F.U., Feijao, A. and S.B. Hedges (2004). A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy and the colonization of land. *BMC Evol. Biol.*, 4: 44-57.

Coenye, T., D. Geveres, Y. Van de Peer, P. Vandamme, and J. Swing (2005). Towards a prokaryotic genomic taxanomy. *FEMS Microbiol. Rev*., 29: 147-167.

Cohan, F.M. (1994a). The effects of rare but promiscuous genetic exchange on evolutionary divergence in prokaryotes. *Am. Natl*., 143: 965-986.

Cohan, F.M. (1994b). Genetic Exchange and evolutionary divergence in prokaryotes. *Trends Ecol. Evol*., 9: 175-180.

Doolittle, R.F. (2005). Evolutionary aspects of whole-genome biology. *Curr. Opin. Struct. Biol*., 15: 248-253.

ElHassan, M. and C. Calladine (1996). Propeller twisting of base pairs and the conformational mobility of dinucleotide steps in DNA. *J. Mol. Biol*., 259: 95-103.

Feil, E.J., E.C. Holmes, D. E. Bessen *et al* (12 co authors). (2001). Recombination within natural populations of pathogenic bacteria: short-term empirical estimates and long-term phylogenetic consequences. *Proc. Natl. Acad. Sci., USA*. 98: 182-187.

Gogarten, J.P., W. F. Doolittle, and J.G. Lawrence (2002). Prokaryotic evolution in light of gene transfer. *Mol. Biol. Evol*., 19: 2226-2238.

Jensen, L. J., C. Friis, and D.W. Ussery (1999). Three views of Microbial Genomes. *Res. Microbiol*., 150:773-777.

Lawrence J.G. and H. Hendrickson (2005). Genome evolution in bacteria: order beneath chaos. *Curr. Opin. Microbiol*., 8: 572-578.

Lawrence, J.G. and H. Ochman (1998). Molecular archaeology of the Escherichia coli genome. *Proc. Natl. Acad. Sci. USA*, 95: 9413-9417.

Levin, B. (1981). Periodic selection, infectious gene exchange, and genetic structure of *E.coli* populations. *Genetics*, 99: 1-23.

May, B.J., Q. Zhang, L.L. Li, M.L. Paustian, T.S. Whittam, and V. Kapur (2001). Complete genomic sequence of *Pasteurella multocida* Pm70. *Proc. Natl. Acad. Sci. USA*, 98: 3460-3465.

Ochman, H., J.G. Lawrence, and E.A. Groisman (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature*, 405: 299-305.

Olson, W., A. Gorin, X. Lu, L. Hock, and V. Zhurkin (1998). DNA sequence dependent deformability deduced from protein-DNA crystal complexes. *Proc. Natl. Acad. Sci. USA*, 95: 11163-11168.

Rhoades, K.R. and R.B. Rimler (1991). *Diseases of poultry*. Iowa State Univ. Press, Ames Publisher, Iowa, USA, pp. 145-162.

Sinden, R., C. Pearson, V. Potaman, and D. Ussery (1998). DNA: Structure and function. *Advances in Genome Biology*, 5A: 1-141.

Worning, P., L.J. Jensen, K.E. Nelson, S. Brunak, and D.W. Ussery (2000). Structural analysis of DNA sequences: Evidence for lateral gene transfer in *Thermotoga maritima*. *Nucleic Acids Res*., 28:706-709.

(*Accepted for publication November 2006*)