IJPSR (2013), Vol. 4, Issue 9 **(Research Article)**

INTERNATIONAL JOURNAL OF PHARMACEUTICAL SCIENCES AND RESEARCH

Received on 29 April, 2013; received in revised form, 15 July, 2013; accepted, 21 August, 2013; published 01 September, 2013

COMPARATIVE STUDY OF OVERLAPPING GENES IN THE GENOMES OF *MYCOPLASMA HOMINIS* AND *MYCOPLASMA PENETRANS*

Kandavelmani Angamuthu*[1] and Shanmughavel Piramanayagam [2]

DBT Bioinformatics Facility, Department of Bioinformatics [1], Research and Development Centre [2], Bharathiar University, Coimbatore – 641046, Tamil Nadu, India

Keywords:

Obligatory parasites, Frame shift, Evolution

Correspondence to Author:

Kandavelmani Angamuthu

Research and Development Centre, DBT Bioinformatics Facility, Department of Bioinformatics, Bharathiar University, Coimbatore – 641046, Tamil Nadu, India

E-mail: kandavelmani@gmail.com

ABSTRACT: Overlapping genes are pairs of adjacent genes whose coding regions are partially or entirely overlapping. Overlapping genes in prokaryotes are a means to minimize genome size and increase the density of genetic information. Overlapping genes play an important role in genome reduction of Mycoplasmas during the course of their degenerative evolution. In this study overlapping genes in the genomes of two obligatory human parasites, *Mycoplasma hominis* and *Mycoplasma penetrans* were extracted and systematically analyzed. Overlapping genes were classified into different categories based on their direction of transcription. Gene pairs that occur as overlapping in both the genomes, gene pairs that occur as overlapping in one genome but are split in the other genome were separated out and more closely analysed. The study revealed that most of the overlapping genes in these genomes are formed due to loss of a stop codon or frame shift. The present study also emphasizes the significance of overlapping genes in the evolution of these genomes.

INTRODUCTION: Overlapping genes are pairs of adjacent genes whose coding regions are partially or entirely overlapping. Overlapping genes are a common occurrence in viruses, bacteria and mitochondria.

Overlapping genes have also been identified in various eukaryotic organisms including humans and systematic analysis methods have been developed [1, 2]. Overlapping genes have evolved due to mutational bias towards deletion [3].

QUICK RESPONSE CODE

DOI: 10.13040/IJPSR.0975-8232.4(9).3504-17

Article can be accessed online on: www.ijpsr.com

DOI link: http://dx.doi.org/10.13040/IJPSR.0975-8232.4(9).3504-17

Evidence from previous studies show that most overlapping genes have originated as a result of stop codon deletion, by a point mutation at the stop codon or by the introduction of a near - end frameshift extending the protein translation till the next in-frame stop codon [4-6].

Overlapping genes are also found to evolve due to the acquisition of an upstream start codon by the downstream gene [7]. Genomic overlaps in prokaryotes may be a result of evolutionary pressure to minimize genome size and increase the density of genetic information [8].

Overlapping genes have been suggested to have multiple functions such as regulation of gene expression, translational coupling and genome overprinting – a process of generating novel genes through accumulated mutations inside a pre – existing gene [9, 10].

Overlapping genes are more conserved between species than non-overlapping genes mostly because a mutation in the overlapping region causes changes in both genes [11, 12]. Overlapping genes can be used as rare genomic markers to get insight into the phylogeny of the completely sequenced microbial genomes [13]. Mycoplasmas, the obligatory parasites of humans, animals and plants seem to have evolved more rapidly than other bacteria since some highly variable positions in their rRNA sequences are strongly conserved in other bacterial species [14].

Mycoplasmas have evolved from Gram-positive bacteria by reductive evolution thereby losing several genes involved in metabolism, cellular process and energy production [15, 16]. However, overlapping genes have been proposed as means of achieving genome reduction by retaining indispensable genes and compressing maximum amount of information in available sequence space [17]. As an endeavor to analyze the role of overlapping genes in genome reduction, in the present study all the overlapping genes in the genomes of the two obligatory human parasites, *Mycoplasma hominis* and *Mycoplasma penetrans* were systematically analyzed.

MATERIALS AND METHODS: The genome sequences of two obligatory human parasites, *M. hominis* and *M. penetrans* were downloaded from the National Center for Biotechnology Information website (ftp://ftp.ncbi.nih.gov/genomes/Bacteria/).

Overlapping genes, defined as pairs of adjacent genes whose coding regions partly or completely overlap were extracted using the (CDS) annotation feature. The extracted overlapping gene pairs were classified into three directional patterns, namely, 'convergent' (→←), 'unidirectional' (→→, ←←), and 'divergent' (←→) based on their direction of transcription [18]. The overlapping gene pairs of *M.penetrans* and *M.hominis* were further classified into four categories,

(i) Gene pairs that occur as overlapping in both genomes

(ii) Gene pairs that overlap in both genomes with different number of overlapping bases

(iii) Gene pair that overlap in *M. penetrans* but are split in *M. hominis*

(iv) Gene pairs that overlap in *M. hominis* but are split in *M. penetrans.*

The gene name, function, direction of overlap and the number of overlapping bases were tabulated for each category an analysed [6].

RESULTS AND DISCUSSION: The number overlapping genes and their orientation in the genomes of *M. hominis* and *M. penetrans* are summarized in **table 1**. Most of the overlapping genes are unidirectional, a few pairs are convergent and only one gene pair in *M. hominis* is divergent. These results concur with the earlier hypothesis that most overlapping gene pairs have unidirectional structure [19]. This is a common occurrence in prokaryote genomes because in prokaryotes adjacent genes with the same orientation are often organized into operons or clusters and are transcribed together in the same direction.

It has been previously reported that the number of overlapping genes increases with the genome size and the number of ORFs [5]. In contrast to this notion, in the current study it is observed that *M. penetrans* with a comparatively larger genome size and higher number of genes is found to have lesser number of overlapping gene pairs than that of *M. hominis* (Table 1). It has been reported earlier that *Mycoplasma genitalium* with the smallest genome has the largest proportion of overlapping gene pairs [6]. This clearly proves that overlapping genes play an important role in genome compaction of the minimal genomes of Mycoplasmas during their course of degenerative evolution.

TABLE 1: NUMBER OF OVERLAPPING GENE PAIRS AND THEIR ORIENTATION

	Genome Size (Mb)	No. of Genes	No. of Overlapping gene pairs	% of Overlapping genes	Orientation of overlapping gene pairs			
					Unidirectional		Convergent	Divergent
					←←	→→	→←	← →
M. hominis	0.67	577	111	19.2	45	61	4	1
M. penetrans	1.36	1069	109	10.1	52	51	6	0

Overlapping genes, number of overlapping nucleotides (length) and their direction of overlap in the genomes of *M. hominis* and *M. penetrans* are listed in **table 2 & 3** respectively. Among all the unidirectional overlapping gene pairs 35 pairs in *M. hominis* and 15 pairs *M. penetrans* overlap only by one base. Most of the overlapping genes in *M. hominis* and *M. penetrans* are formed due to loss of a stop codon or frame shift.

Table 4 summarizes the gene pairs that are found to be overlapping in both the genomes. All the 12 common overlapping gene pairs are found to have unidirectional orientation.

Among the 12 common overlapping gene pairs, five gene pairs overlap by the same number of nucleotides.

Whereas, seven other common overlaps differ in their overlapping lengths. The five common unidirectional overlapping gene pairs are found on the same strand in both the genomes except for MHO_1230: MHO_1240 (→→) and MYPE7420: MYPE7430 (←←). There are 36 gene pairs that are overlapping in *M. hominis* but are non-overlapping in *M. penetrans* (**Table 5**). Whereas on the other hand, there are 17 gene pairs that are overlapping in *M. penetrans* but are split in *M. hominis* which has a smaller genome (**Table 6**). These discrepancies are in accordance with the notion that overlapping genes may be a means of compressing maximum amount of information into the available short sequence space may be a result of evolutionary pressure to minimize the genome size and increase the density of genetic information [20].

TABLE 2: OVERLAPPING GENES IN *M. HOMINIS*

S. No.	Genes	Genes	Length	Direction
1	MHO_0010	rnpA	26	←←
2	dnaN	MHO_0060	1	→→
3	ksgA	MHO_0090	8	←←
4	MHO_0090	MHO_0100	1	←←
5	MHO_0190	atpB	8	→→
6	atpA	atpG	22	→→
7	atpD	atpC	1	→→
8	scpA	scpB	11	→→
9	scpB	rluC	14	→→
10	gatC	gatA	11	→→
11	gatA	gatB	8	→→
12	gatB	MHO_0460	8	→→
13	MHO_0580	MHO_0590	1	→→
14	cysS	MHO_0620	1	→→
15	MHO_0660	fba	17	←←
16	MHO_0740	MHO_0750	1	→→
17	MHO_0750	MHO_0760	59	→→
18	MHO_0840	uvrC	11	→→
19	rplk	rplA	1	→→

20	MHO_0960	recR	1	→→
21	dnaH	MHO_1000	8	→→
22	thrS	trpS	1	←←
23	MHO_1070	MHO_1080	8	→→
24	MHO_1080	MHO_1090	50	→→
25	nusB	tlyA	17	→→
26	tlyA	nifS	1	→→
27	nifS	nifU	14	→→
28	nifU	mucB	8	→→
29	mucB	MHO_1260	1	→→
30	alaS	MHO_1320	13	→→
31	MHO_1320	MHO_1330	8	→→
32	MHO_1330	MHO_1340	11	→→
33	MHO_1340	greA	1	→→
34	MHO_1400	rluD	2	← →
35	oppD	oppF	8	→→
36	MHO_1570	MHO_1580	1	→→
37	prfA	hemK	1	→→
38	cmk	engA	14	→→
39	MHO_1710	MHO_1720	14	←←
40	oppB	oppC	17	→→
41	tktA	MHO_1780	11	←←
42	ruvA	ruvB	26	→→
43	hisS	aspS	17	→→
44	dgk	hpt	13	→→
45	cbiO1	cbiO2	10	→→
46	cbiO2	cbiQ	8	→→
47	ligA	MHO_2010	24	→←
48	MHO_2010	MHO_2020	76	←←
49	MHO_2060	pgsA	11	→→
50	smf	nfo	9	→←
51	MHO_2190	MHO_2200	8	←←

52	MHO_2330	MHO_2340	20	←←
53	MHO_2370	ribF	1	→→
54	MHO_2450	MHO_2460	1	←←
55	infA	map	1	←←
56	map	adk	10	←←
57	adk	secY	7	←←
58	secY	rplO	1	←←
59	rpsE	rplR	1	←←
60	rplN	rpsQ	8	←←
61	rpsQ	rpmC	1	←←
62	rplP	rpsC	23	←←
63	rpsC	rplV	1	←←
64	rpsS	rplB	1	←←
65	rplW	rplD	1	←←
66	rplD	rplC	1	←←
67	lspA	ileS	11	←←
68	MHO_3080	MHO_3090	14	→→
69	atpD	atpA	1	←←
70	MHO_3140	MHO_3150	8	←←
71	MHO_3150	MHO_3160	23	←←
72	MHO_3170	MHO_3180	20	←←
73	hsdS	hsdS	23	←←
74	MHO_3280	MHO_3290	1	←←
75	rnhB	MHO_3310	8	←←
76	MHO_3360	MHO_3370	14	→←
77	uvrA	MHO_3430	11	←←
78	MHO_3440	MHO_3450	38	←←
79	gidB	prs	26	←←
80	ldh	pgi	8	←←
81	pgi	rpsA	11	←←
82	pepF	apt	1	←←
83	md2	md1	8	←←

84	rpII	MHO_3930	20	←←
85	engB	gcp	1	←←
86	gcp	MHO_3960	7	←←
87	MHO_3960	MHO_3970	1	←←
88	MHO_4000	rpmG	1	←←
89	dnaE	polA	14	→→
90	recD	pth	1	→→
91	pth	tilS	14	→→
92	MHO_4320	ung	11	→←
93	MHO_4380	gltX	20	←←
94	potA	potB	32	→→
95	potB	potC	8	→→
96	MHO_4500	MHO_4510	7	→→
97	pyrH	frr	1	→→
98	MHO_4660	MHO_4670	1	→→
99	plsX	rnc	11	→→
100	MHO_4750	MHO_4760	10	→→
101	MHO_4760	MHO_4770	26	→→
102	MHO_4870	sps1	8	→→
103	MHO_4950	bcrA	1	→→
104	bcrA	MHO_4970	23	→→
105	MHO_4970	MHO_4980	8	→→
106	MHO_4980	MHO_4990	23	→→
107	nusA	MHO_5260	7	→→
108	MHO_5260	infB	49	→→
109	infB	rbfA	14	→→
110	MHO_5310	MHO_5320	8	→→
111	MHO_5330	MHO_5340	1	←←

TABLE 3: OVERLAPPING GENES IN *M. PENETRANS*

S. No.	Genes	Genes	Length	Direction
1	spoU	MYPE240	10	→→
2	MYPE340	MYPE350	32	→→

3	MYPE380	scpB	11	→→
4	MYPE400	engB	1	→→
5	MYPE430	MYPE440	109	→→
6	MYPE450	MYPE460	23	→→
7	MYPE490	fmt	40	→→
8	prfA	hemK	35	→→
9	MYPE670	MYPE680	35	→→
10	hrcA	MYPE720	26	→→
11	nusA	MYPE1100	8	→→
12	infB	rbfA	17	→→
13	nfo	fur	13	→→
14	MYPE1280	MYPE1290	17	→→
15	MYPE1290	MYPE1300	26	→←
16	MYPE1300	rluB	11	←←
17	Lsp	MYPE1330	8	←←
18	MYPE1440	pth	8	→→
19	MYPE1460	MYPE1470	7	→→
20	mucB	MYPE1500	11	←←
21	Efp	MYPE1590	1	→→
22	MYPE1590	nusB	11	→→
23	MYPE1610	MYPE1620	11	→→
24	MYPE1620	MYPE1630	38	→→
25	Dam	dam	38	→→
26	MYPE1830	ligA	23	→→
27	MYPE1970	dnaE	23	→→
28	Fpg	MYPE2010	28	→→
29	MYPE2020	MYPE2030	44	→→
30	gatA	gatB	23	→→
31	MYPE2200	MYPE2210	8	←←
32	engA	gpsA	11	→→
33	MYPE2390	MYPE2400	53	→→
34	MYPE2470	MYPE2480	20	→→

35	ruvB	ruvA	35	←←
36	MYPE2830	MYPE2840	4	→←
37	hprK	lgt	14	→→
38	MYPE3040	rpe	4	→←
39	MYPE3070	MYPE3080	62	→→
40	mraW	MYPE3230	1	→→
41	MYPE3350	dgkA	8	→→
42	Era	MYPE3380	1	→→
43	glyS	dnaG	11	→→
44	plsX	rnc	26	→→
45	tpiA	pgm	11	→→
46	MYPE4190	MYPE4200	55	→←
47	ulaA	tktA	1	→→
48	tktA	MYPE4520	14	→→
49	MYPE4520	MYPE4530	20	→→
50	MYPE4600	MYPE4610	1	→→
51	MYPE4640	MYPE4650	8	→→
52	MYPE4730	MYPE4740	17	→→
53	MYPE4850	MYPE4860	16	→→
54	MYPE4930	MYPE4940	20	←←
55	hemN	miaA	14	←←
56	pdhA	pdhB	1	→→
57	oppC	oppB	8	←←
58	MYPE5580	MYPE5590	8	←←
59	MYPE5590	MYPE5600	29	←←
60	MYPE5620	MYPE5630	14	←←
61	MYPE5630	gmk	1	←←
62	MYPE5680	MYPE5690	17	←←
63	glpQ	MYPE5760	8	←←
64	Htp	MYPE5820	1	←←
65	MYPE5870	MYPE5880	1	←←
66	MYPE5880	MYPE5890	1	←←

67	MYPE5890	MYPE5900	35	←←
68	MYPE6010	MYPE6020	17	←←
69	MYPE6320	MYPE6330	17	←←
70	MYPE6440	MYPE6450	20	←←
71	MYPE6495	MYPE6500	35	←←
72	MYPE6600	MYPE6610	11	←←
73	MYPE6610	MYPE6620	8	←←
74	dhfR	thyA	1	←←
75	thyA	folD	14	←←
76	araD	sgaU	20	←←
77	sgaU	ulaD	26	←←
78	MYPE7375	MYPE7380	53	←←
79	nifU	nifS	14	←←
80	MYPE7570	MYPE7580	38	←←
81	MYPE7680	MYPE7690	16	←←
82	fruK	fruA	47	←←
83	fruA	MYPE7770	11	←←
84	pyrB	pyrR	1	←←
85	gidB	MYPE8000	14	→→
86	MYPE8090	MYPE8100	8	←←
87	hsdR	hsdR	50	←←
88	MYPE8350	sigA	20	←←
89	MYPE8510	MYPE8520	16	←←
90	MYPE8530	alaS	17	←←
91	potA	potB	32	→→
92	potB	potC	8	→→
93	potC	MYPE8600	13	→→
94	MYPE8600	gcp	23	→→
95	MYPE8620	MYPE8630	17	→→
96	MYPE8730	oppF	8	←←
97	MYPE9020	cysS	8	←←
98	truB	MYPE9290	4	→←

99	metS	MYPE9390	8	←←
100	MYPE9420	MYPE9430	4	→←
101	MYPE9440	MYPE9450	13	←←
102	MYPE9640	MYPE9650	11	←←
103	cbiO	cbiO	16	←←
104	MYPE9870	MYPE9880	8	←←
105	MYPE9880	MYPE9890	28	←←
106	rpmC	rplP	14	←←
107	rplW	rplD	1	←←
108	rplD	rplC	1	←←
109	MYPE10380	ksgA	20	←←

TABLE 4: OVERLAPPING GENE PAIRS COMMON TO BOTH *M. HOMINIS* AND *M. PENETRANS*

S. No.	Gene Name	Function	Locus		Length of Overlap		Direction of Overlap	
			M.hom	*M.pen*	*M.hom*	*M.pen*	*M.hom*	*M.pen*
1	prfA: hemK	peptide chain release factor 1: S-adenosylmethionine-dependent methyltransferase	MHO_1600: MHO_1610	MYPE650: MYPE660	1	35	→ →	→ →
2	infB: rbfA	translation initiation factor IF-2: ribosome-binding factor A	MHO_5270: MHO_5280	MYPE1110: MYPE1120	14	17	→ →	→ →
3	gatA: gatB	aspartyl/glutamyl-tRNA amidotransferase subunit A: aspartyl/glutamyl-tRNA amidotransferase subunit B	MHO_0440: MHO_0450	MYPE2100: MYPE2110	8	23	→ →	→ →
4	ruvB: ruvA	Holliday junction DNA helicase B: holliday junction DNA helicase	MHO_1850: MHO_1860	MYPE2780: MYPE2790	26	35	→ →	← ←
5	plsX: rnc	putative glycerol-3-phosphate acyltransferase PlsX: ribonuclease III	MHO_4680: MHO_4690	MYPE3640: MYPE3650	11	26	→ →	→ →
6	oppC: oppB	oligopeptide transport system permease protein oligopeptide transport system permease protein	MHO_1740: MHO_1750	MYPE5540: MYPE5550	17	8	→ →	← ←
7	nifU: nifS	nitrogen fixation protein NifU: aminotransferase NifS	MHO_1230: MHO_1240	MYPE7420: MYPE7430	14	14	→ →	← ←
8	PotA: potB	spermidine/putrescine transport ATP-binding protein: spermidine/putrescine transport system permease	MHO_4450: MHO_4460	MYPE8570: MYPE8580	32	32	→ →	→ →
9	potB: potC	spermidine/putrescine transport system permease: spermidine/putrescine transport system permease Po	MHO_4460: MHO_4470	MYPE8580: MYPE8590	8	8	→ →	→ →

10	cbiO: cbiO	cobalt transporter ATP-binding subunit: cobalt transporter ATP-binding subunit	MHO_1970: MHO_1980	MYPE9760: MYPE9770	10	16	→ →	← ←
11	rplW: rplD	50S ribosomal proteinL23: 50S ribosomal protein L4	MHO_2960: MHO_2970	MYPE1010: MYPE1010	1	1	← ←	← ←
12	rplD: rplC	50S ribosomal protein L4: 50S ribosomal protein L3	MHO_2970: MHO_2980	MYPE10170: MYPE1018	1	1	← ←	← ←

TABLE 5: GENE PAIRS THAT OVERLAP IN *M. HOMINIS* BUT ARE SPLIT IN *M. PENETRANS*

S. No.	Gene name	Function	Locus	Length	Direction
1	atpA: atpG	ATP synthase subunit alpha: ATP synthase subunit gamma	MHO_0240: MHO_0250	22	→ →
2	atpD: atpC	ATP synthase subunit beta: ATP synthase subunit epsilon	MHO_0260: MHO_0270	1	→ →
3	scpA: scpB	segregation and condensation protein A: segregation and condensation protein B	MHO_0400: MHO_0410	11	→ →
4	scpB: rluC	segregation and condensation protein B: ribosomal large subunit pseudouridine synthase	MHO_0410: MHO_0420	14	→ →
5	gatC: gatA	glutamyl- tRNA(gln)amidotransferase subunit C: glutamyl-tRNA(gln)amidotransferase subunit A	MHO_0420: MHO_0440	11	→ →
6	thrS: trpS	threonyl-tRNA synthetase: Tryptophanyl-tRNA synthetase	MHO_1050: MHO_1060	1	← ←
7	nusB: tlyA	N utilization NusB-like protein: Hemolysin A	MHO_1210: MHO_1220	17	→ →
8	tlyA: nifS	Hemolysin A: Nitrogen fixation protein NifS(aminotransferase	MHO_1220: MHO_1230	1	→ →
9	oppD: oppF	putative oligopeptide transport ATP-binding protein: putative oligopeptide transport ATP-binding protein	MHO_1540: MHO_1550	8	→ →
10	Cmk: engA	Cytidylate kinase: GTP-binding protein engA	MHO_1670: MHO_1680	14	→ →
11	hisS: aspS	Histidyl-tRNA synthetase: Aspartyl-tRNA synthetase	MHO_1870: MHO_1880	17	→ →
12	Dgk: hpt	Deoxyguanosine kinase: Hypoxanthine-guanine phosphoribosyltransferase (HGP...	MHO_1930: MHO_1940	18	→ →
13	Smf: nfo	DNA processing protein smf: Endonuclease IV	MHO_2130: MHO_2140	9	→ ←
14	infA: map	translation initiation factor IF-1: methionine aminopeptidase (MAP)	MHO_2760: MHO_2770	1	← ←
15	Map: adk	methionine aminopeptidase (MAP): Adenylate kinase (ATP-AMP transphosphorylase)	MHO_2770: MHO_2780	10	← ←
16	Adk: secY	Adenylate kinase (ATP-AMP transphosphorylase): Preprotein translocase secY subunit	MHO_2780: MHO_2790	7	← ←

17	secY: rplO	Preprotein translocase secY subunit: 50S ribosomal protein L15	MHO_2790: MHO_2800	1	← ←
18	rplN: rpsQ	50S ribosomal protein L14: 30S ribosomal protein S17	MHO_2880: MHO_2890	8	← ←
19	rpsQ: rpmC	30S ribosomal protein S17: 50S ribosomal protein L29	MHO_2890: MHO_2900	1	← ←
20	rplP: rpsC	50S ribosomal protein L16: 30S ribosomal protein S3	MHO_2910: MHO_2920	23	← ←
21	rpsC: rplV	30S ribosomal protein S3: 50S ribosomal protein L22	MHO_2920: MHO_2930	1	← ←
22	rpsS: rplB	30S ribosomal protein S19: 50S ribosomal protein L2	MHO_2940: MHO_2950	1	← ←
23	lspA: ileS	lipoprotein signal peptidase: Isoleucyl-tRNA synthetase	MHO_3040: MHO_3050	11	← ←
24	atpD: atpA	ATP synthase subunit beta: ATP synthase subunit alpha	MHO_3120: MHO_3130	1	← ←
25	hsdS: hsdS	Type I restriction enzyme specificity protein: Type I restriction enzyme specificity protein	MHO_3220: MHO_3230	23	← ←
26	gidB: prs	methyltransferase gidB: ribose-phosphate pyrophosphokinase	MHO_3540: MHO_3550	26	← ←
27	Ldh: pgi	L-lactate dehydrogenase: glucose-6-phosphate isomerase	MHO_3580: MHO_3590	8	← ←
28	Pgi: rpsA	glucose-6-phosphate isomerase: 30S ribosomal protein S1	MHO_3590: MHO_3600	11	← ←
29	pepF: apt	Oligoendopeptidase F: Adenine phosphoribosyl transferase	MHO_3670: MHO_3680	1	← ←
30	md2: md1	ABC transporter ATP binding protein: ABC transporter ATP binding protein	MHO_3820: MHO_3830	8	← ←
31	engB: gcp	putative GTP-binding protein engB: O-sialoglycoprotein endopeptidase	MHO_3940: MHO_3950	1	← ←
32	DnaE: polA	DNA polymerase III subunit alpha: DNA polymerase I	MHO_4120: MHO_4130	14	→ →
33	recD: pth	Exodeoxyribonuclease V subunit alpha: Peptidyl-tRNA hydrolase	MHO_4200: MHO_4210	1	→ →
34	Pth: tilS	Peptidyl-tRNA hydrolase: tRNA(Ile)-lysidine synthase	MHO_4210: MHO_4220	14	→ →
35	pyrH: frr	Uridylate kinase smbA: ribosome recycling factor	MHO_4600: MHO_4610	1	→ →
36	infB: rbfA	translation initiation factor IF-2: ribosome-binding factor A	MHO_5270: MHO_5280	14	→ →

TABLE 6: GENE PAIRS THAT OVERLAP IN *M. PENETRANS* BUT ARE SPLIT IN *M. HOMINIS*

S. No.	Gene name	Function	Locus	Length	Direction
1	nfo: fur	endonuclease IV: ferric uptake regulation protein	MYPE1190: MYPE1200	13	→ →
2	Dam: dam	adenine-specific DNA methyltransferase: adenine-specific DNA methyltransferase	MYPE1780: MYPE1790	38	→ →
3	engA: gpsA	GTP-binding protein EngA: NAD-dependent glycerol-3-phosphate dehydrogenase	MYPE2290: MYPE2300	11	→ →
4	hprK: lgt	HPr kinase/phosphorylase: prolipoprotein diacylglyceryl transferase	MYPE2990: MYPE3000	14	→ →
5	glyS: dnaG	glycyl-tRNA synthetase: DNA primase	MYPE3390: MYPE3400	11	→ →
6	tpiA: pgm	triose phosphate isomerase [Mycoplasma penetrans H: phosphoglyceromutase	MYPE3730: MYPE3740	11	→ →
7	ulaA: tktA	ascorbate-specific PTS system enzyme IIC: transketolase	MYPE4500: MYPE4510	1	→ →
8	hemN: miaA	coproporphyrinogen III oxidase : tRNA-isopentenyl pyrophosphate transferase	MYPE5010: MYPE5020	14	← ←
9	pdhA: pdhB	Pyruvate dehydrogenase E1 component subunit alpha: pyruvate dehydrogenase E1 component subunit beta	MYPE5080: MYPE5090	1	→ →
10	dhfR: thyA	dihyrofolatereductase: thymidylate synthase	MYPE6860: MYPE6870	1	← ←
11	thyA: folD	thymidylate synthase: methylenetetrahydrofolate dehydrogenase	MYPE6870: MYPE6880	14	← ←
12	araD: sgaU	L-ribulose-5-phosphate 4-epimerase: L-xylulose 5-phosphate 3-epimerase	MYPE7160: MYPE7170	20	← ←
13	sgaU: ulaD	L-xylulose 5-phosphate 3-epimerase: 3-keto-L-gulonate-6-phosphate decarboxylase	MYPE7170: MYPE7180	26	← ←
14	fruK: fruA	1-phosphofructokinase: PTS system fructose-specific IIABC component	MYPE7750: MYPE7760	47	← ←
15	pyrB: pyrR	aspartate carbamoyltransferase catalytic subunit: pyrimidine regulatory protein PyrR	MYPE7890: MYPE7900	1	← ←
16	hsdR: hsdR	type I restriction-modification system R subunit: type I restriction-modification system R subunit.	MYPE8220: MYPE8230	50	← ←
17	rpmC: rplP	ribosomal protein L29: 50S ribosomal protein L16	MYPE10100: MYPE10110	14	← ←

CONCLUSION: Computational comparative genomic analysis of *M. hominis* and *M. penetrans* has provided a breakthrough towards the understanding of the evolution of these genomes. It has been analyzed that there is no correlation between the genome size, number of genes and number of overlapping gene pairs. Most of the overlapping genes in *M. hominis* and *M. penetrans* are formed due to loss of a stop codon or frame shift. The current study reveals the significance of overlapping genes in bringing about the genome compaction of Mycoplasmas during the course of reductive evolution.

REFERENCES:

1. Ho MR, Tsai KW and Lin WC: A unified framework of overlapping genes: towards the origination and endogenic regulation. Genomics 2012; 100(4):231-9.

2. Kim DS, Cho CY, Huh JW, Kim HS and Cho HG EVOG: a database for evolutionary analysis of overlapping genes. Nucleic Acids Research 2009; 37: D698-702.
3. Clark MA, Baumann L, Thao ML, Moran NA, and Baumann P: Degenerative minimalism in the genome of a psyllid endosymbiont. Journal of Bacteriology 2001; 183: 1853–1861.
4. Fukuda Y, Washio T, and Tomita M: Comparative study of overlapping genes in the genomes of Mycoplasma genitalium and Mycoplasma pneumoniae. Nucleic Acids Research 1999; 27: 1847–1853.
5. Fukuda Y, Nakayama Y, and Tomita M: On dynamics of overlapping genes in bacterial genomes. Gene 2003; 323:181–187.
6. Sakharkar KR, Sakharkar MK, Verma C and Chow VT: Comparative study of overlapping genes in bacteria, with special reference to Rickettsia prowazekii and Rickettsia conorii. International Journal of Systematic and Evolutionary Microbiology 2005; 55:1205–1209.
7. Cock PJ and Whitworth DE: Evolution of relative reading frame bias in unidirectional prokaryotic gene overlaps. Molecular Biology and Evolution 2010; 27: 753-756.
8. Lillo F and Krakauer DC: A statistical analysis of the three- fold evolution of genomic compression through frame overlaps in prokaryotes. Biology Direct 2007; 2:22.
9. Keese PK and Gibbs A: Origins of genes: ‘‘big bang’’ or continuous creation? Proceedings of the National Academy of Sciences of the United States of America 1992; 89: 9489–9493.
10. Johnson ZI and Chisholm SW: Properties of overlapping genes are conserved across microbial genomes. Genome Research 2004; 14: 2268–2272.
11. Lipman DJ: Making (anti)sense of non-coding sequence conservation. Nucleic Acids Research 1997; 25: 3580–3583.
12. Yelin R, Dahary D, Sorek R, *et al*: Widespread occurrence of antisense transcription in the human genome. Nature Biotechnology 2003; 21:379–386.
13. Luo Y, Fu C, Zhang DY and Lin K: Overlapping genes as rare genomic markers: the phylogeny of gamma-Proteobacteria as a case study. Trends in Genetics 2006; 22: 593-596.
14. Weisburg WG, Tully JG, Rose DL, Petzel JP, Oyaizu H, Yang D, Mandelco L, Sechrest J, Lawrence TG, Etten VJ, Maniloff J and Woese CR: A Phylogenetic analysis of mycoplasmas: basis for their classification. Journal of Bacteriology 1989; 171: 6455-6467.
15. Fadiel A, Eichenbaum KD, El Semary N, and Epperson B: Mycoplasma genomics: tailoring the genome for minimal life requirements through reductive evolution. Frontiers in Bioscience. 2007; 12:2020-8.
16. Razin S and Hayflick L: Highlights of mycoplasma research-An historical perspective. Biologicals 2010; 38: 183-190.
17. Sakharkar KR and Chow VT: Strategies for genome reduction in microbial genomes. Genome Informatics 2005; 16: 69-75.
18. Rogozin IB, Spiridonov AN, Sorokin AV, Wolf YI, Jordan IK, Tatusov RL, and Koonin EV: Purifying and directional selection in overlapping prokaryotic genes. Trends in Genetics 2002; 18: 228–232.
19. Eyre-Walker A: The distance between Escherichia coli genes is related to gene expression levels. Journal of Bacteriol 1995; 177: 5368–5369.
20. Krakauer DC: Stability and evolution of overlapping genes. International Journal of Systematic and Evolutionary Microbiology 2000; 54: 731–739.

How to cite this article:
Angamuthu K and Piramanayagam S: Comparative study of overlapping genes in the genomes of *Mycoplasma hominis* and *Mycoplasma penetrans*. *Int J Pharm Sci Res* 2013: 4(9); 3504-3517. doi: 10.13040/IJPSR. 0975-8232.4(9).3504-17