

Conservation and metabolic functional significance of overlapping gene in the bacterial genomes

Phatthanaphong Wanchanthuek^{1,3}, David J. Hampson² and Matthew Bellgard^{1*}

¹Centre for Comparative Genomics, Murdoch University, South Street, Murdoch, Western Australia 6150, Australia

²School of Veterinary and Biomedical Sciences, Murdoch University, South Street, Murdoch, Western Australia 6150, Australia

³Faculty of Informatics, Mahasarakham University, Kantarawichai District, Mahasarakham 44150, Thailand

ABSTRACT

Overlapping genes, which are common in prokaryotes, are adjacent chromosomally-positioned genes that share overlapping DNA in their respective open reading frame. Previous studies have suggested that overlapping genes are a result of evolution mechanisms toward genome reduction as well as functional mechanisms for gene coexpression, transcription efficiency and translational coupling. To date, little has been reported on the possible association of overlapping gene in the same metabolic pathways. This finding would imply that overlapping genes are not only likely the result of functional constraints but also constrained from a metabolomic context. With the recently in-house sequenced genomes of two *Brachyspira* sequences and the availability of 10 other available spirochete genomes an in-depth comparative genomics analyses has been conducted to investigate overlapping gene. Interestingly, the number of overlapping in the 12 spirochete genome ranges from 11% (in *B. hyodysenteriae*) and 45% (in *T. palidum*). Of these overlapping genes 80% were unidirectional ($\rightarrow\rightarrow/\leftarrow\leftarrow$). A small number of overlapping gene-pairs were duplicated within each genome and there were some triplets overlapping gene. Orthologous overlapping genes were identified within the genus. For example, 19 orthologous overlapping genes were shared between *B. hyodysenteriae* and *B. pilosicoli*. In a similar way, 11, 38 and 36 of orthologous overlapping genes were shared in *Treponema*, *Borrelia* and *Leptospira*, respectively. These within-genus orthologous overlapping gene-pairs were unique and were not shared with other spirochete genus.

Utilizing currently available metabolic pathway information, it was observed that over 75% of overlapping genes identified were in the same or related metabolic pathway. For example, eight overlapping gene-pairs in *Brachyspira* genomes were found in four metabolic pathways related to the purine metabolism pathway such as folate metabolism, glutamate metabolism, oxidative phosphorylation and glyoxylate. Of the 25% remaining of overlapping genes, 50% contained one hypothetical gene with unknown function. In addition, in one of the orthologous overlapping genes in *Brachyspira* species a promoter is shared indicating a special class of overlapping gene operon. We discuss these finding in context with evolution, function and habitat of spirochetes. These resulted suggest that overlapping gene are likely result of metabolomic constraints acting on the species.

Keywords: Coexpression of overlapping gene in metabolic pathway

*Corresponding Author:

Professor Matthew Bellgard, PhD.

Centre for Comparative Genomics

Murdoch University, South Street, Murdoch, Western Australia 6150, Australia

Phone: + 61 (8) 9360 6088 Fax: + 61 (8) 9360 7238

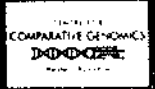
E-mail: mbellgard@ccg.murdoch.edu.au

Submitted to the 20th Annual Meeting and International Conference of the Thai Society for Biotechnology at the Taksila Hotel, Maha Sarakham, Thailand on October 14-17th, 2008.

Conservation and metabolic functional significance of overlapping gene in the bacterial genomes



Phatthanaphong Wanchanthuek^{1,3}, David J. Hampson² and Matthew Bellgard¹



¹Centre for Comparative Genomics, Murdoch University, South Street, Murdoch, Western Australia 6150, Australia
²School of Veterinary and Biomedical Sciences, Murdoch University, South Street, Murdoch, Western Australia 6150, Australia
³Faculty of Informatics, Maharakham University, Kantarawickai District, Maharakham 44150, Thailand

Introduction

Overlapping genes, which are common in prokaryotes, are adjacent chromosomally-positioned genes that share overlapping DNA in their respective open reading frame. Previous studies have suggested that overlapping genes are a result of evolution mechanisms toward genome reduction as well as functional mechanisms for gene co-expression, transcription efficiency and translational coupling. To date, little has been reported on the possible association of overlapping gene in the same metabolic pathways. This finding would imply that overlapping genes are not only likely the result of functional constraints but also constrained from a metabolomic context. With the recently in-house sequenced genomes of two *Brachyspira* sequences and the availability of nine other available spirochaete genomes an in-depth comparative genomics analyses has been conducted to investigate overlapping gene.

Methods

Nine spirochaete genomes are publicly available: three species of *Borrelia*, four species of *Leptospira*, two species of *Treponema* including *E. coli*. These genome sequences were downloaded from the NCBI. The complete genomic sequences of *B. hyodysenteriae* WA-1 and a draft of genome sequenced of *B. pilosicoli* 95/1000 were recently sequenced. The annotation of the sequence was carried out at the Centre for Comparative Genomics (CCG), Murdoch University.

Overlapping genes were extracted from each genome using Perl according to the annotations (Fig. 1). Three pattern of overlapping genes are identified: unidirectional ($\rightarrow\rightarrow$ or $\leftarrow\leftarrow$), divergent ($\rightarrow\leftarrow$) convergent ($\leftarrow\rightarrow$) (1,2).

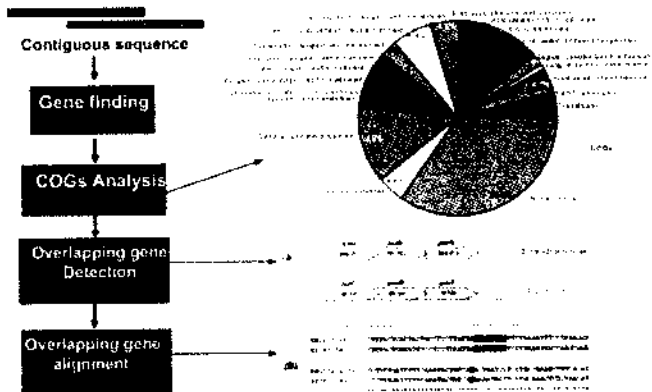


Figure 1. Overlapping gene analysis in bacterial genomes

Results

The majority of overlapping genes overlap are only a few base pairs (1-4 bp) (2). More than 80% of overlapping genes in spirochaete genomes were unidirectional. Overlapping genes can be orthologous only within genus (Fig. 2) (4). Many of the functional categories of orthologous overlapping gene that are involved in essential housekeeping functions, such as DNA and RNA metabolism, protein processing and secretion, cell structure, cellular processes, and energetic and intermediary metabolism, are represented in the core gene set.

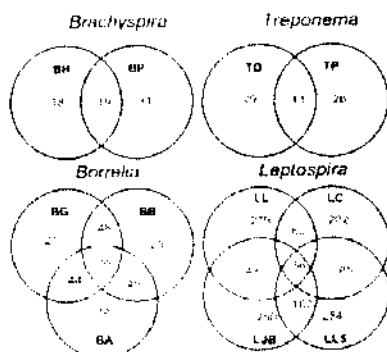


Figure 2. Venn diagram showing the number of orthologous of overlapping gene-pairs in spirochaete genomes

>75% of overlapping gene-pairs among spirochaete genome were unidirectionally transcribed (1) and assigned to the same or related metabolic pathways in the KEGG database (Fig. 3). This result showed that co-expression of overlapping gene-pairs is caused by clustering of overlapping of genes in the same and related metabolic pathway. In *Brachyspira*, many overlapping gene-pairs are co-regulated in the same and related metabolic pathways as a single entity. An interesting example of this type of co-regulation can be found in purine biosynthesis pathway (Fig. 4). Moreover, eight overlapping gene-pairs were found in four metabolic pathways related to the purine metabolism pathway such as folate metabolism, glutamate metabolism, oxidative phosphorylation and glyoxylate.

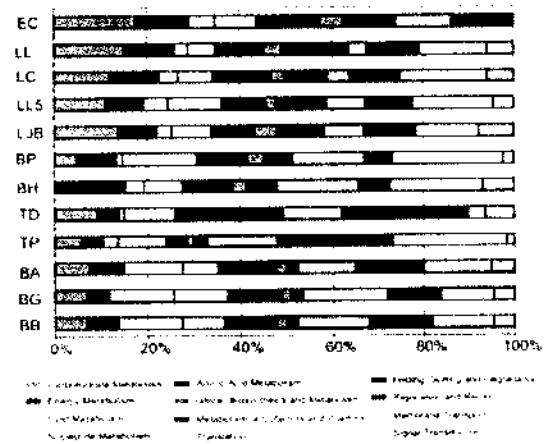


Figure 3. Distribution of overlapping genes pair of the inferred genome-wide in KEGG metabolic pathway

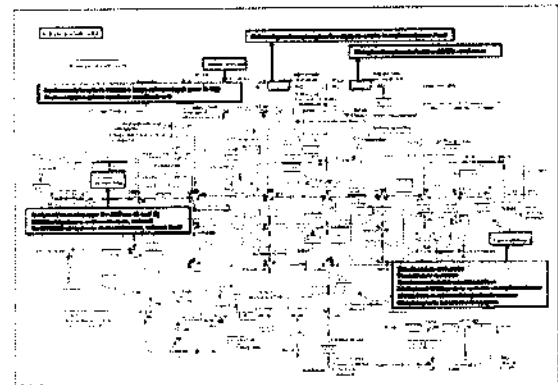


Figure 4. Overlapping gene-pair involved in purine biosynthesis of *B. hyodysenteriae* and *B. pilosicoli*

Conclusion

Acknowledgements

The Centre for Comparative Genomics is supported by the Government of Western Australia Centre of Excellence in Science and Innovation Program. PW would like to thank the people at the Centre for Comparative Genomics for their contributions. PW thanks to the Royal Thai Government for supporting his PhD study and the Faculty of Informatics, Maharakham University, Thailand.

References

1. Fukuda, Y., Yoshida, T. and Tamaki, M. 1999. Comparative study of overlapping genes in the genomes of *Chlamydia pneumoniae* and *Chlamydia trachomatis*. *Microb. Drug Resist.* 17:1847-1853.
2. Fukuda, Y., Nakamura, Y., and Tamaki, M. 2003. Co-expression of overlapping genes in bacterial genomes. *Gene* 323:181-189.
3. Nishikubo, M., Sugawara, G., Lator, D., Fuku, J., Fuku, P. and Lator, U. 2001. An alternative form of equal and metabolic pathway. *Genetics* 157:105-107.
4. Johnson, ZI and Chikalin, S.V. 2004. Pangenom of overlapping genes at bacterial mass: metabolic genome. *Genome Research* 14:268-277.