

An investigation of conserved coexpression in bacteria

Nels Thorsteinson

Research and Training Centre on Bioinformatics
Institute for Information Transmission Problems



Russian Academy of Sciences



Биоинформатика



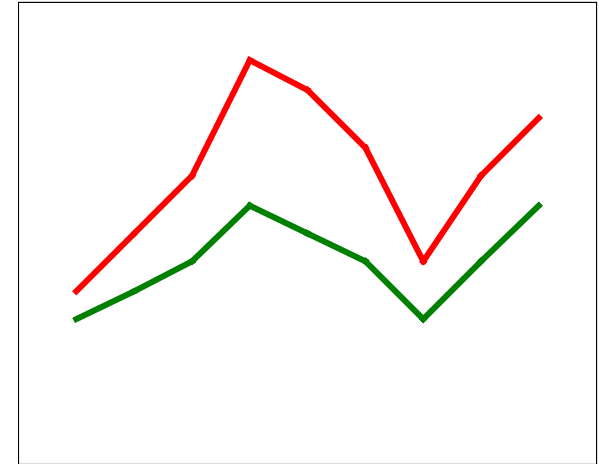
CIHR/MSFHR Strategic Training Program in

BIOINFORMATICS

Introduction

- Coexpression

- groups of genes with similar expression profiles
- measured by Pearson correlation
- involved in similar functions



- Conserved coexpression

- groups of genes which are coexpressed in multiple species
- involved in core biological processes

Methods

- Public data from GEO, Array Express, Stanford

Escherischia coli

Bacillus subtilis

Mycobacterium tuberculosis

Vibrio cholera

Streptococcus pneumonia

Campylobacter jejuni

Streptomyces coelicor

- NCBI's COG database
 - Orthologue assignment
- STRING database
 - Evaluation of coexpression networks

Figure 1: Similarity of single genome coexpression sets

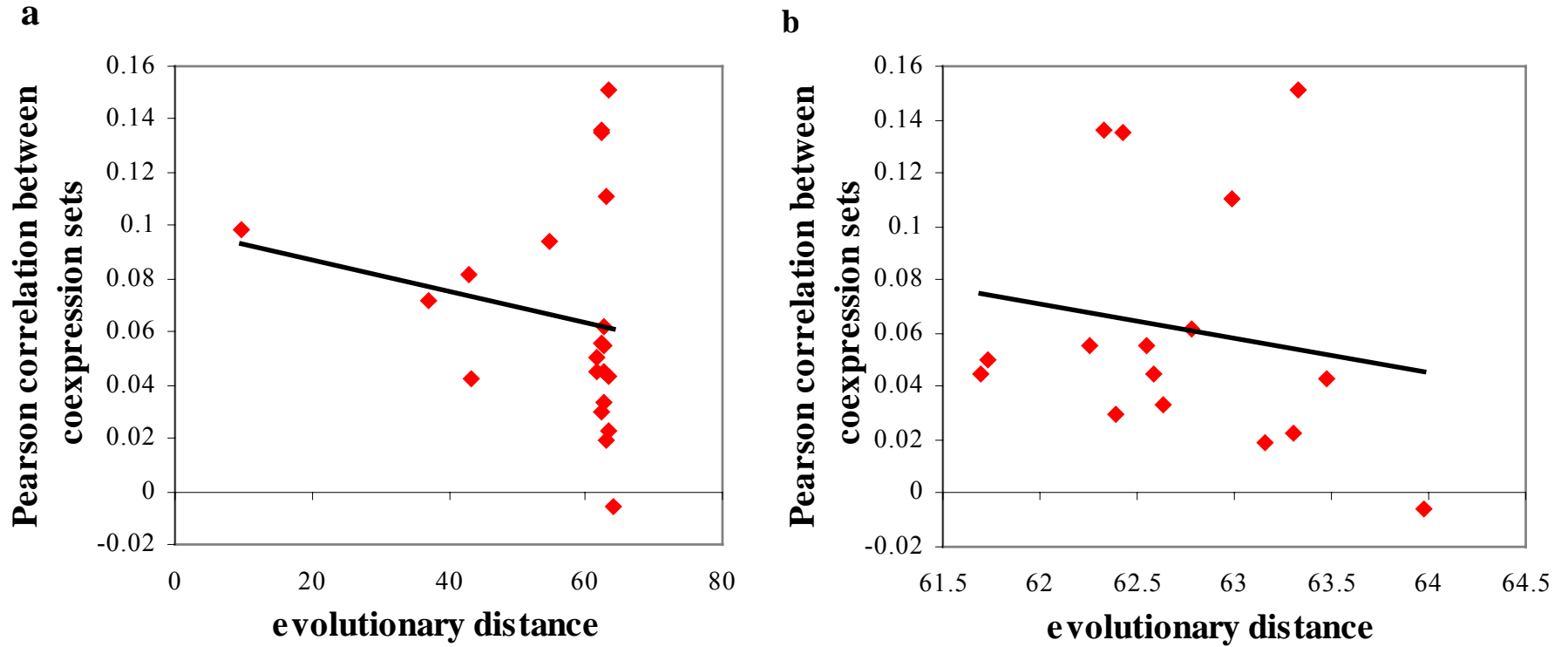


Figure 2: Correlation of coexpression sets to STRING's neighbourhood score

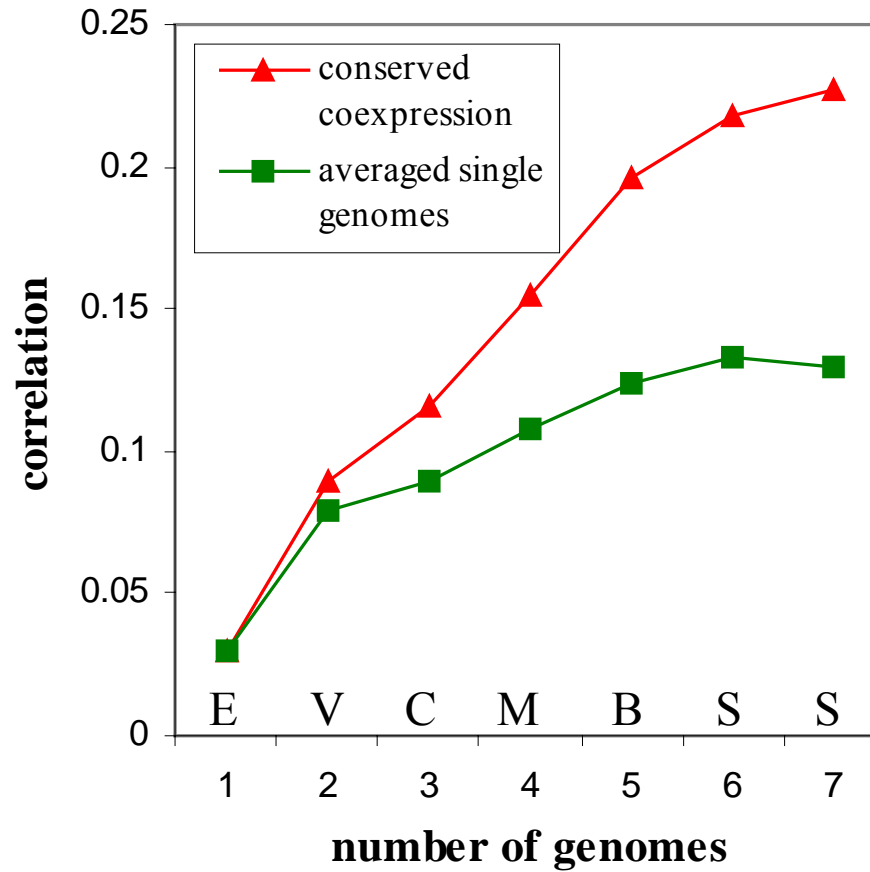
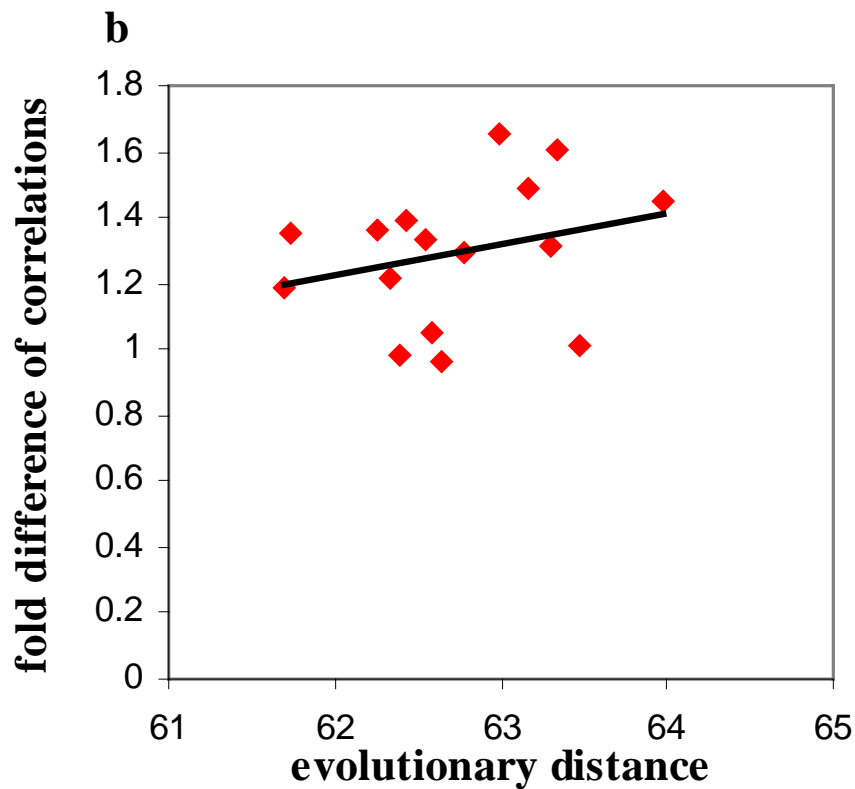
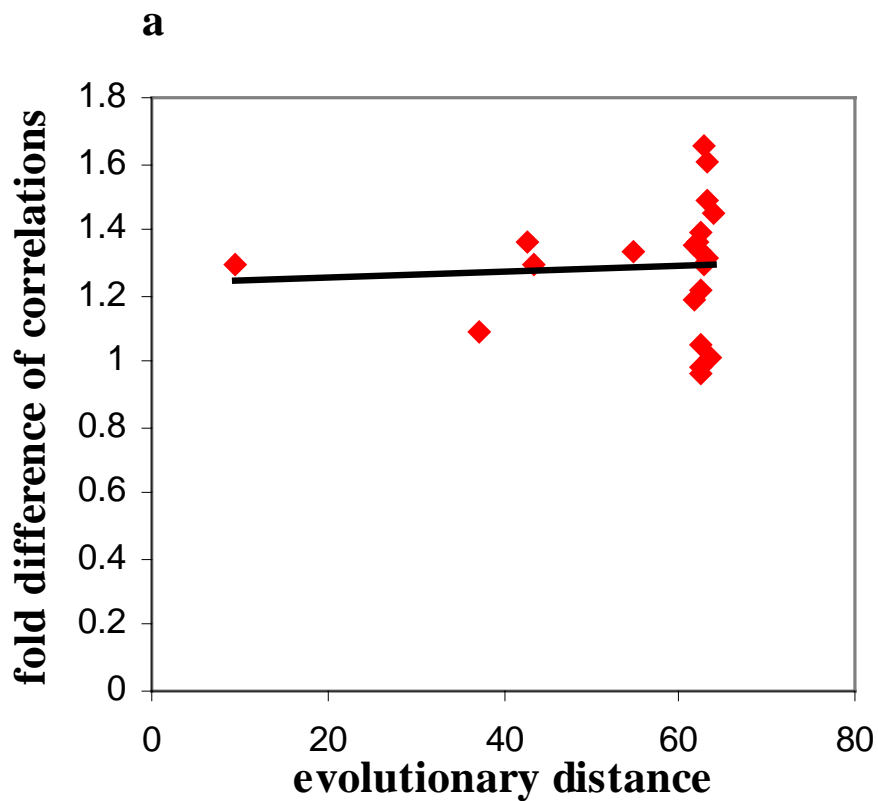


Figure 3: Correlation of conserved coexpression sets to STRING's neighbourhood score



Methods

- Functional classification of the genes in the conserved coexpression network

Functional Classification	Number of genes
Translation, ribosomal	47
Energy production	9
Transcription	7

- Only one third of gene pairs consist of genes belonging to the same operon

Conclusion

- The more genomes used when calculating a conserved coexpression network, the higher the correlation to functional interactions
- The further the distance between the species for which a conserved coexpression network is calculated, the higher the correlation of the resulting network to functional interactions
- Presented conserved coexpression network

Acknowledgements

Mikhail Gelfand

Anya Gerasimova

Alexey Kazakov

Artem Cherkasov

Research and Training Centre on Bioinformatics
Institute for Information Transmission Problems



Russian Academy of Sciences



Биоинформатика



CIHR/MSFHR Strategic Training Program in

BIOINFORMATICS