Identification of Clusters of Orthologous Groups of proteins that correlate with sodium-dependent bioenergetics

Ilya M. Bushmakin[1], Olesya I. Klimchuk[1], Armen Y. Mulkidjanian[1,2,3], Daria V. Dibrova[2]

[1]Faculty of Bioengineering and Bioinformatics, Lomonosov Moscow State University; [2]Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University; [3]Department of Physics, Osnabrueck University

Introduction

The majority of living organisms use proton-motive force as an intermediate in their membrane bioenergetics processes, however, some prokaryotes utilizes the sodium motive force instead. The sodium bioenergetics is shown to be the primordial type of bioenergetics [1, 2]. Here we have analyzed prokaryotic genomes with a goal to idetify those membrane protein coding genes that correlate with sodium-dependent bioenergetics. Finding proteins associated with the type of bioenergetics might help clarify to pinpoint the core, presumably ancestral set of proteins related to generation of sodium gradient.

Materials

- We have used Clusters of Orthologous Groups (COGs) database and correlation analysis to reveal proteins associated with bioenergetics.
- Presence of Na^+-binding ATPase is used as a proxy of sodium bioenergetics.
- Original COG DB contains 711 prokaryotic genomes and 4632 COGs. After excluding organisms with both sodium and proton ATPases present in genomes 606 organisms remained.
- Predicted via TMHMM non-membrane COGs (COGs with less than 10% proteins containing 2 or more helices) were removed from the dataset.
- Very common COGs (60% and more occupancy) were also excluded. Only 710 COGs remained in the final matrix.

Methods

- Obtained matrix was used to find correlations between COGs and usage of Na+ as a coupling ion.
- The Dice correlation was chosen as a method of finding correlation coefficients [3].
- COGs with coefficients more than 0.167 were considered further, the chance of randomly obtaining greater coefficients is 0.05.
- Homologs of well-correlated COGs were found via COGcollator [4] and PSI-BLAST. Genome surroundings were examined with COGNAT [5].

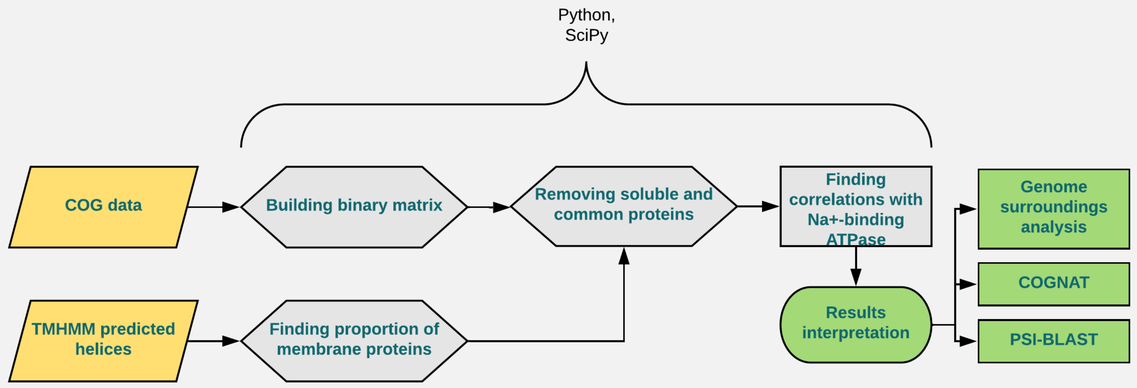

Figure 1: Methodics overview

Results

91 COGs has correlation > 0.167, including RNF subunits and Na^+-transporting decarboxylase which are known to be major sources of sodium ion gradient. The found COGs split into six major groups:

- Na^+-dependent transporters
- ABC-type transporters
- Complex I homologs (Eha, Mnh, Formate hydrogenlyase)
- Sodium-dependent tetrahydromethanopterin S-methyltransferase subunits
- Proteins involved in lipopolysaccharide synthesis
- Completely uncharacterised COGs

As seen from Tab. 1, the best correlated COGs mostly contain uncharacterized proteins. We could attribute some of the uncharacterized proteins to sodium transporters based on weak sequence similarity. Some deeply examined COGs are shown in the Tab. 2.

Table 1: COGs with the highest correlation with sodium-dependent bioenergetics

COG	Dice correlation
COG1563 Uncharacterized MnhB-related	0.467
COG1269 V-type H+-ATPase subunit I/STV1	0.395
COG1827 Transcript. regulator of NAD metabolism	0.391
COG4945 Carbohydrate-binding DOMON domain	0.366
COG2426 Uncharacterized membrane protein	0.347
COG1906 Uncharacterized protein	0.347
COG4769 Uncharacterized membrane protein	0.337
COG1822 Uncharacterized membrane protein	0.333
COG4720 Uncharacterized membrane protein	0.333
COG1852 Uncharacterized protein	0.315
COG2456 Uncharacterized protein	0.308
COG1883 Na+-transporting decarboxylase, beta s/u	0.303
COG4708 Uncharacterized membrane protein	0.292
COG1750 Predicted archaeal protease, S18 family	0.289
COG4035 Uncharacterized membrane protein	0.286
COG4042 Energy-converting hydrogenase Eha s/u A	0.286
COG1784 TctA family transporter	0.283
COG2034 Uncharacterized membrane protein	0.282
COG4039 Energy-converting hydrogenase Eha s/u C	0.276

Table 2: Examined COGs

COG	Dice correlation	Homologous cogs	Surroundings
COG4769	0.337	EcfT	Fe-S associated ABC transporter
COG4720	0.333	EcfT	Fe-S associated ABC transporter
COG2456	0.308	GT2 glycosyltransferase	GT2-family proteins
COG2426	0.350	Small multidrug efflux pump	–
COG4078	0.271	NuoM/L/N	Mnh and Eha subunits
COG0650	0.256	NuoH	–
COG1006	0.177	NuoK	Mnh subunits, O-antigen ligase

Conclusions

- Export of fermentations products in symport with sodium ions, as postulated earlier [6] might be one of main sources of sodium-motive force
- Na^+ antiporters may convert other gradients to Na^+ gradient, which can be used to produce ATP
- Lipopolysaccharide synthesis might depend on sodium-motive force, which probably gives energy to carbohydrate transporters
- Many homologs of Complex I appear to translocate Na+ ions

References

[1] A. Y. Mulkidjanian, M. Y. Galperin, K. S. Makarova, Y. I. Wolf, and E. V. Koonin, "Evolutionary primacy of sodium bioenergetics," *Biology Direct*, vol. 3, no. 1, p. 13, 2008.

[2] A. Y. Mulkidjanian, P. Dibrov, and M. Y. Galperin, "The past and present of sodium energetics: May the sodium-motive force be with you," *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, vol. 1777, no. 7-8, pp. 985–992, 2008.

[3] L. R. Dice, "Measures of the Amount of Ecologic Association Between Species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.

[4] D. V. Dibrova, K. A. Konovalov, V. V. Perekhvatov, K. V. Skulachev, and A. Y. Mulkidjanian, "COGcollator: a web server for analysis of distant relationships between homologous protein families," *Biology Direct*, vol. 12, no. 1, p. 29, 2017.

[5] O. I. Klimchuk, K. A. Konovalov, V. V. Perekhvatov, K. V. Skulachev, D. V. Dibrova, and A. Y. Mulkidjanian, "COGNAT: a web server for comparative analysis of genomic neighborhoods," *Biology Direct*, vol. 12, no. 1, p. 26, 2017.

[6] T. A. Michel and J. M. Macy, "Generation of a membrane potential by sodium-dependent succinate efflux in selenomonas ruminantium.," *Journal of Bacteriology*, vol. 172, no. 3, pp. 1430–1435, 1990.

Contact Information

- Email: isnicsi@gmail.com