

Genes	Genome									
	YP	YE	PM	AA	HI	HD	VV	VP	VC	VF
<i>atpIBEFHAGDC</i>	^{1a} <i>atpIBEFHAGDC</i>		^{1b} <i>atpBEFHAGDC</i>				^{1a} <i>atpIBEFHAGDC</i>			
<i>napFDAGHBC</i>	^{2a} <i>napFDABC</i>		^{2b} <i>napFDAGHBC</i>				^{2a} <i>napFDABC</i> ^{2c} <i>napGH</i>		^{2a} <i>napFDABC</i>	
<i>nrfABCDECFG</i>	0	0	^{3a} <i>nrfABCDECFG</i>		^{3b} <i>nrfABCD</i> ^{3c} <i>nrfEXFG</i>		^{3d} <i>nrfA</i> ^{3e} <i>nrfBCDECF</i> ^{3f} <i>nrfG</i>		0	^{3d} <i>nrfA</i> ^{3f} <i>nrfG</i>
<i>torCAD</i>	0	0	^{4a} <i>torCAD</i>	0	0	0	^{4b} <i>torCA</i> ^{4c} <i>torD</i>			
<i>moaABCDE</i>	^{5a} <i>moaACDE</i>		^{5a} <i>moaACDE</i>				^{5b} <i>moaABCDE</i>			
<i>pdhR-aceEF-lpdA</i>	^{6a} <i>pdhR-aceEF-lpdA</i>		^{6b} <i>aceEF-lpdA</i>				^{6a} <i>pdhR-aceEF-lpdA</i>			
<i>sucABCD</i>	^{7a} <i>sucABCD</i>		^{7a} <i>sucABCD</i>		^{7b} <i>sucAB</i> ^{7c} <i>sucCD</i>	^{7a} <i>sucABCD</i>	^{7a} <i>sucABCD</i>			
<i>nagBACD</i>	^{8a} <i>nagBACD</i>		^{8b} <i>nagBA</i> ^{8c} <i>nagC</i>		^{8b} <i>nagBA</i>		^{8d} <i>nagB</i> ^{8e} <i>nagAC</i>			
<i>deoCABD</i>	^{9a} <i>deoCABD</i>		^{9b} <i>deoD</i>			^{9c} <i>deoCD</i>	^{9a} <i>deoCABD</i>			
<i>malQ-glgBXCAP</i>	^{10a} <i>malQ</i> ^{10b} <i>glgBXCAP</i>		^{10c} <i>malQ-glgBXCAP</i>			0	^{10d} <i>malQ</i> ^{10e} <i>glgB</i> ^{10f} <i>glgX</i> ^{10g} <i>glgC</i> ^{10h} <i>glgA</i>			
<i>acpP-fabF</i>	^{11a} <i>acpP-fabF</i>		^{11b} <i>acpP</i>				^{11a} <i>acpP-fabF</i>			
<i>narQP</i>	^{12a} <i>narP</i>		^{12a} <i>narP</i> ^{12b} <i>narQ</i>				^{12c} <i>narQP</i>			

Description of additional data files

Ravcheev_Gerasimova_Add1.pdf

PDF file.

Title of data: Phylogenetic trees for FdnG/FdoG (a), FdnH/FdoH (b) and FdnI/FdoI.

Description of data: The trees were constructed by the neighbour-joining method. The expected fraction of amino acid substitutions is indicated for branches longer than 0.05. Genome abbreviations: EC – *Escherichia coli*, ST – *Salmonella typhi*, EO – *Erwinia carotovora*, YP – *Yersinia pestis*, YE – *Y. enterocolitica*, PM – *Pasteurella multocida*, AA – *Actinobacillus actinomycetemcomitans*, HI – *Haemophilus influenzae*.

The analysis of the trees indicates that the most parsimonious evolutionary scenario is the duplication of the fumarate dehydrogenase operon in the ancestral Enterobacteria and further species-specific loss of the Fdn copy in the *Yersinia* lineage. In any case, it is clear that the fumarate dehydrogenase of the *Yersinia* spp. is Fdo.

Ravcheev_Gerasimova_Add2.pdf

PDF file.

Title of data: Additional predicted regulatory interactions

Description of data: “For genome abbreviations see “Materials and Methods”. Candidate sites are shown by letters: F – Fnr sites, A – ArcA sites, N – NarP sites; conserved sites are shown by capital letters, non-conserved ones, by lower-case letters (for details see “Materials and Methods”). Absence of the corresponding sites is shown by dashes. Absent genes are shown by zeros. Superscripts point to changes in operon structures described in detail in additional file 3. #: operon forms a divergon with another one.”

Ravcheev_Gerasimova_Add3.pdf

PDF file.

Title of data: Additional observed changes in the operon structures

Description of data: For genome abbreviations see “Materials and Methods”. Superscripts coincide with the corresponding superscripts in additional file 2.

Ravcheev_Gerasimova_Add4.pdf

PDF file.

Title of data: False-positive predictions

Description of data: The last column lists the reasons to exclude the operons. For genome abbreviations see “Materials and Methods”.

Ravcheev_Gerasimova_Add5.pdf

PDF file.

Title of data: Phylogenetic trees for NarL/NarP (a) and NarX/NarQ (b).

Description of data: The trees were constructed by the neighbour-joining method. The expected fraction of amino acid substitutions is indicated for branches. The branches corresponding to *Yersinia* spp. proteins are shown with broken lines. Genome abbreviations: EC – *Escherichia coli*, ST – *Salmonella typhi*, EO – *Erwinia carotovora*, YP – *Yersinia pestis*, YE – *Y. enterocolitica*, PM – *Pasteurella multocida*, AA – *Actinobacillus actinomycetemcomitans*, HI – *Haemophilus influenzae*, HD – *Haemophilus ducreyi*, VV – *Vibrio vulnificus*, VP – *V. parahaemolyticus*, VC – *V. cholerae*, VF – *Vibrio fischeri*.

Ravcheev_Gerasimova_Add6.pdf

PDF file.

Title of data: Position weight matrices (profiles) for Fnr (a), ArcA (b) and NarP (c) binding sites

Description of data: Rows: nucleotides. Columns: Positions. Cells: positional nucleotide weights.

Ravcheev_Gerasimova_Add7.pdf

PDF file.

Title of data: Sequence logos for the Fnr (a), ArcA (b) and NarP (c) binding sites

Description of data: Horizontal axis, position in the binding site; vertical axis, information content in bits. The height of each column is proportional to the positional information content in the given position; the height of each individual symbol reflects its prevalence in the given position.