

## Supporting Information

We enclose this File S1 of Supporting Information containing a detailed account of the classification of bacterial *COX* operons (2 pages), 4 additional Figures and 4 additional Tables.

### Classification of bacterial *COX* operons

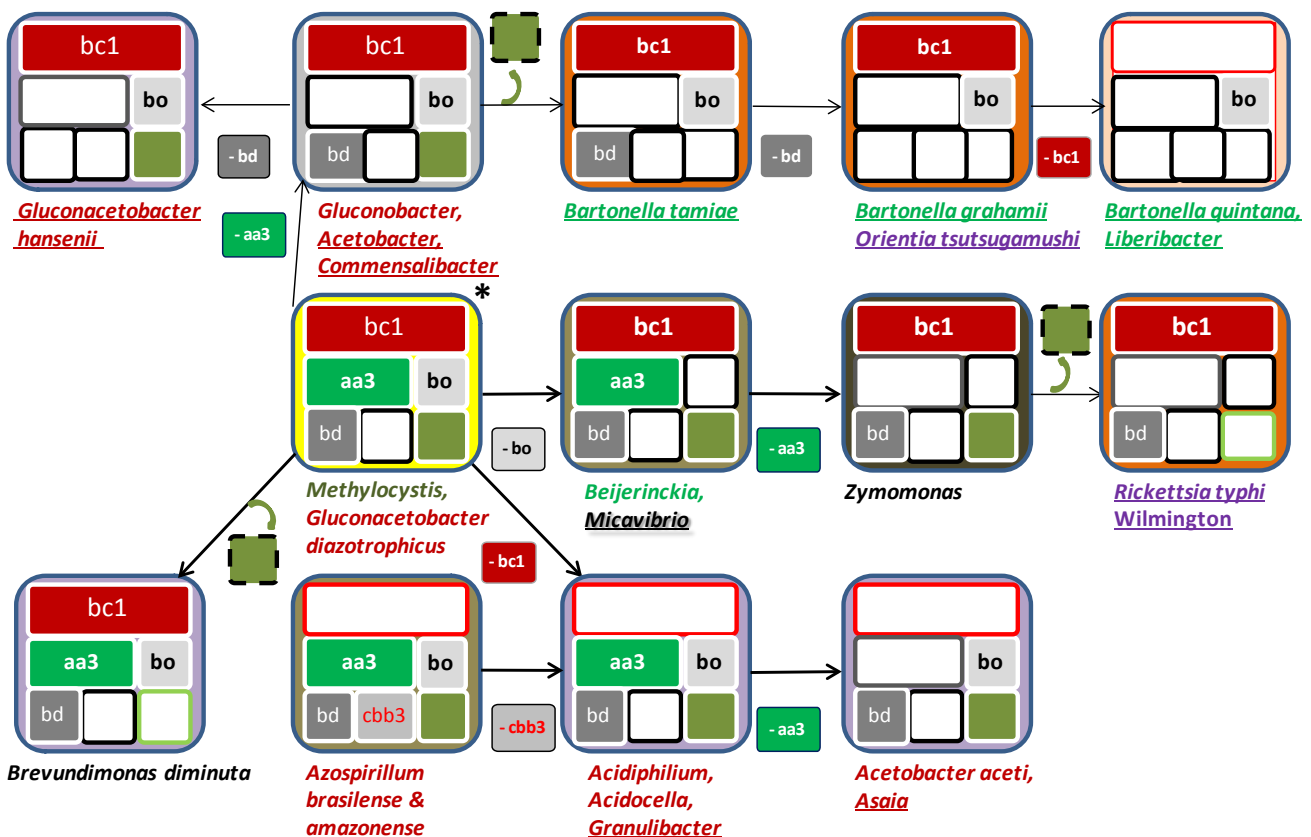
The terminal oxidases considered here (Fig. 1) belong to the super-class of heme-copper oxygen reductases, which are classified on the basis of structural-functional features of their common catalytic subunit [26], corresponding to *COX1* or *ctaD* for the *aa<sub>3</sub>* cytochrome *c* oxidases. Type A includes the majority of these oxidases, while type B includes the *ba<sub>3</sub>* oxidases [26,47] that lack *COX3* and *COX4* in their core gene sequence. Type C correspond to the *cbb<sub>3</sub>* oxidases, which also lack *COX3* and *COX4*. Finally, NOR includes the nitric oxide reductases, which are considered under the system of N metabolism here (Fig. 1). While maintaining the framework of this classification, we introduce a new subdivision of the *aa<sub>3</sub>* cytochrome *c* oxidases which accounts for the large variety in their genomic sequences organized in operons [46,47,51], herein called ‘*COX* operons’ (Figs. 3 and S4, Table S2). We have now classified them systematically in three groups (or sub-types of the overall category A [26]): **type a**, conserving the core sequence *ctaC-F*; **type b**, with the core sequence intermixed with genes for the assembly of the oxidase cofactors, e.g. *CtaB*; and **a-b transition**, showing features intermediate between type a and b and restricted to  $\beta$ - and  $\gamma$ - proteobacteria. Additionally, the *ba<sub>3</sub>* oxidases (type B [26]) are maintained as a separated type of *COX* operon (Table S2), even though their gene sequences shows features related to both *COX* operon a-b transition and the operon of *cbb<sub>3</sub>* cytochrome *c* oxidases. From the composition and positioning of the genes that are prepended [82] (i.e. upstream) and appended [82] (i.e. downstream) to the core sequence, we distinguish four different *COX* operons of type a. Type **a** alone is characterized by *COX1* fused with *COX3* and the presence of tri-heme *c* cytochromes (Fig. 3 and data not shown). Type **a-I** is associated with the Act (Alternative Complex Three) gene clusters containing homologues of the nitrite reductase *nrfD* [49,50] (Figs. 3A and S4); similarly to *ba<sub>3</sub>* oxidases, the combined operon is predominantly present in Rhizobiales and Rhodospirillales (Fig. 6A). Type **a-II** is recognized for the fusion of *COX4* at the C-terminus of *COX1* (Fig. S2), the fusion of *COX2* with a *c*

cytochrome and the frequent association with PQQ-dependent dehydrogenases; its richness in mono- and di-heme *c* cytochromes suggests possible evolutionary relationships with the operon of *cbb<sub>3</sub>* oxidases. Finally, type **a-III** characteristically contains a doublet of *ctaE/COX3* proteins (Figs. 3A and S4) and is widely distributed among  $\alpha$ -proteobacteria (Table S2). Although *COX* operons of type a are present also in  $\beta$ - and  $\gamma$ -proteobacteria (Tables S2-S4), *COX* operon type **b** is present only in  $\alpha$ -proteobacteria, showing no fusion among *COX* subunits and more variation in the contiguity of the gene sequence than in the gene sequence itself. Indeed, while *Beijerinckia* or *Tistrella* have a continuous gene sequence, Rhodobacterales and other taxa present different fragmentations of the operon in separate gene clusters (Figs. 3A and S4A). Of note, previous phylogenetic studies on *aa<sub>3</sub>* cytochrome *c* oxidases predominantly used the proteins belonging to *COX* operon type b [6,7,13,47,64].

With regard to the possible evolution of *COX* operons, we hypothesize that an early duplication of *COX4* might have been crucial in the progressive breaking apart of the operon core sequence and the major structural transition of *COX3* from the 5-helices to the 7-helices form (Fig. S3), presumably by fusion of a two-helices *COX4*-like protein. This large molecular change has enhanced *COX3* binding to specific lipids, in particular PG1 and PG2 [60] (see also Table S4). The resulting dense lipo-protein packing could modulate the entry of oxygen in the catalytic centre [60], thereby reducing the affinity for the oxygen substrate. This molecular change affecting the affinity for oxygen appears to correspond to the juncture between *COX* operons type a-II lacking a separate *COX4* subunit and the transition a-b type (Figs. 3 and S3). We surmise that such a junction might have been concurrent with the separation of the  $\beta$ - and  $\gamma$ -lineages from primordial  $\alpha$ -proteobacteria (Fig. 3C), an event which occurred around the time when oxygen levels strongly increased on the planet [31].

## Supporting Information Figures – and their legend


**Figure S1 - Pathways for the bioenergetic evolution of  $\alpha$  bacterial not leading to mitochondria.** The diagram shows the additional subsets of bioenergetic systems that are not shown in Fig. 1B, including those of *Asaia* and *Saccharibacter* (Table S1B). The asterisk\* labels the same subset as in Fig. 1B (main text), but with fewer representative taxa. Underlined organisms are symbionts or pathogens. Each of the six bioenergetic systems presented in Fig. 1 was identified from its catalytic protein subunits and was considered functionally absent when one or more of these subunits were not found in their completeness, as indicated by the profile of their conserved domains (cf. [41]). The functional absence of a given system is represented by an empty square as in Fig. 1B.

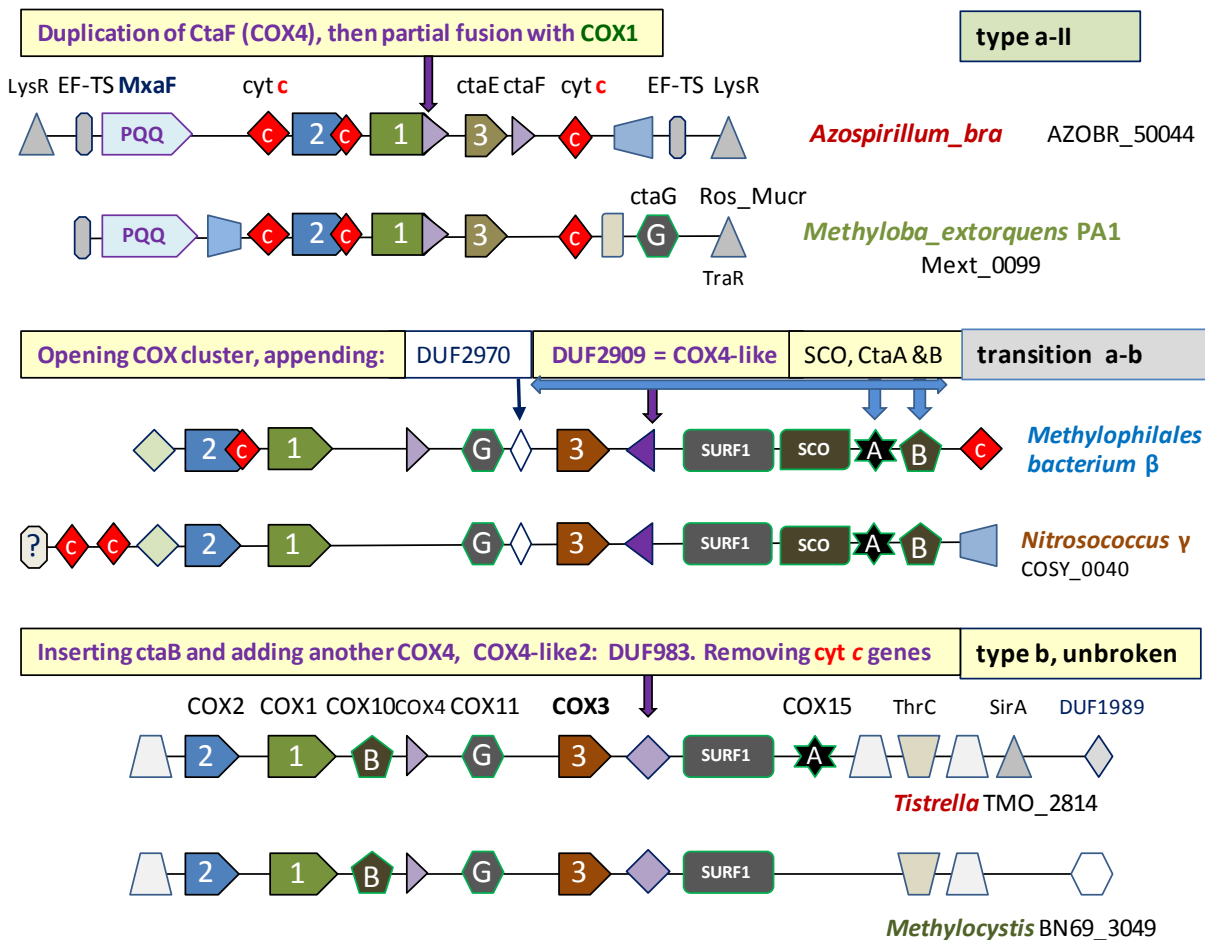


**Fig. S2. Sequence analysis to identify the fusion of COX4 subunit with COX1 proteins.**

**A.** Sequences of recognised or putative COX4 were manually aligned to reference proteins having known 3D structure around the first transmembrane helix (TM1, highlighted in grey): subunit IV of *Thermus caa<sub>3</sub>* oxidase (accession: pdb|2YEV [54]) and subunit IV (COX4\_pro\_2 super family [cl06738]) of *Rhodobacter Sphaeroides aa<sub>3</sub>* oxidase (chain D, accession: pdb|1M57 [53]). \*Residues in **bold** have positive scores ( $\geq 0$ ) in the BLOSUM62 substitution matrix [83], those **yellow-highlighted** are identical with either reference protein, while those highlighted in **purple** are identical to *Janibacter COXIV* (accession: ZP\_00994995) with scores  $\geq 5$  [83]. The total count of identities is also highlighted in yellow (**tot**) before the description of the protein on the right. It was used to identify other COX4-like proteins such as DUF983 (see Fig. 3A and the section entitled “classification of bacterial COX operons”). The minimal count for deeming a protein as “COX4-like” was considered to be 10, but several COX1 proteins exhibited larger numbers of identities. The region of ciliate COX1 showing similarity with COX4 partially overlaps the last transmembrane region (TM12) of aligned COX1, which is well conserved among all available COX1 sequences from ciliates. However, the COX4-like region in bacterial COX1 and that of the pathogenic fungus *Zygomycetozia* [55] lies outside the conserved domains of other COX1 proteins. *Azospirillum bras*, *Azospirillum brasilense*; *Methylobac\_extor*, *Methylobacterium extorquens*. **B.** This panel shows the alignment of COX4 subunits around the second transmembrane helix (TM2), the structure of which is known only for subunit IV of *Thermus caa<sub>3</sub>* [54] that was used as the reference for aligning bacterial COX4 and mtDNA-encoded proteins. In **bold black** are the residues that are identical in the aligned position of at least two COX4 sequences, or are positive substitutions [83] across at least three aligned COX4 sequences; they are additionally **yellow-highlighted** when identical between at least one bacterial COX4 and one mtDNA-encoded protein (cf. A). In **bold dark blue** are the residues that are positive substitutions between bacterial COX4 and mtDNA-encoded proteins, while those in **bold light blue** are identical or positive substitutions among the aligned mtDNA-encoded proteins. This colour labelling enhances the limited similarity between the sequences shown.



**Fig. S3. Gene sequence of additional COX operons in diverse bacteria.** The reference gene name for each cluster is indicated on the right of the figure. Symbols identify the same proteins as in Fig. 3A, with the addition of , protein related to nucleotide exchange factor EF-TS. These short proteins were recognised after alignment to the sequence with known 3D structure of Chain A, dimerization domain of Ef-Ts From *Thermus thermophilus* (Accession: pdb|1TFE|A) using a sequence analysis similar to that shown in Fig. S2. Hypothetical steps in the evolution of COX operons are indicated.



**Figure S4. Structure-based alignment of bacterial and mitochondrial “Rieske” ISP.**

The protein sequences of various ISP of the *bc<sub>1</sub>* complex were aligned following structures available from various sources matching the alignment gaps or insertions with the most refined 3D data [68-71]. The limits of secondary structures (alpha helices, highlighted in green, and beta sheets, highlighted in purple) were deduced from a consensus of the latest coordinates deposited in the NCBI databanks [68-71]. Common insertions and deletions (Indels [72]) between mitochondrial and bacterial sequences are consecutively labelled CIMit1-7 (cf.

Fig. 7A). The C terminus of some sequences is truncated at the residue indicated by the numeral before the slash. Key residues for the iron-sulfur cluster, including Y165 influencing its redox potential [71], are in **bold**. Note that *Nitrospira*, *Nitrosomonas*, *Nitrosococcus* and *Methylocystis* are metabolically related by ammonia/methane autotrophy. The organisms follow established phylogenetic distance from top to bottom according to the following taxonomic groups and species. **Cyanobacteria:** *Synechocystis* (*b<sub>6</sub>f* complex), *Synechocystis* sp. PCC 6803, 192 aa; **Nitrospirales:** *Nitrospira*, *Candidatus Nitrospira defluvii* [73], 183 aa ; **ε-proteobacteria:** *Epsilon*, *Helicobacter pylori*, 167 aa; **Planctomycetes:** *Kuenenia\_2*, *Candidatus Kuenenia stuttgartiensis* (in-operon Kuste3096 [66]), 173 aa ; **Schlesneria\_2**, *Schlesneira paludicola* DSM 18645 (accession: ZP\_11092182), 189 aa. **γ-proteobacteria:** *Nitrosoc*, *Nitrosococcus watsonii* C-113, 201 aa; **Frateuria**, *Frateuria aurantia*, 201 aa; **β-proteobacteria:** *Nitrosomonas*, *Nitrosomonas europaea* ATCC 19718, 201 aa; **Beta**, *Neisseria meningitidis* MC58, 193 aa. **α-proteobacteria:** *Methylocy\_1 &\_2*, *Methylocystis* sp. SC2 [84], *\_1* in-operon, 176 aa , *\_2* in isolated gene cluster, 209 aa; *Methylob\_r*, *Methylobacterium radiotolerans* JCM 2831, 189 aa; **Nitrobacter**, *Nitrobacter hamburgensis* ISP2, 219 aa; **Gluc\_dia**, *Gluconacetobacter diazotrophicus* PAI 5 (in isolated gene cluster), 221 aa; **Saccharib**, *Saccharibacter* sp. (Chouaia *et al* [22]), 223 aa; *Glu\_oxyd*, *Gluconobacter oxydans* H24, 218 aa; **Beijerinckia**, *Beijerinckia indica*, 172 aa; **RoseobacterA2**, *Roseobacter litoralis petA2* in-operon, 186 aa; **Maricaulis\_1**, *Maricaulis maris* in-operon, 207 aa; **Micavibrio**, *Micavibrio aeruginosavorus* [25], 185 aa; **Magnetococcus**, *Magnetococcus marinus* [76], 178 aa; **Rickettsia**, *Rickettsia felis*, 177 aa. **Mitochondria:** **Acanthamoeba**, *Acanthamoeba castellanii*, 235 aa; **S\_cerevisiae**, *Saccharomyces cerevisiae*, mature 185 aa (3D structure available [85]); **Chicken**, *Gallus gallus*, mature 192 aa (3D structure available [68]). C-terminal extensions are highlighted in pale blue with some conserved residues in gray.

Rieske ISP, N-terminal part: UCR TM domain, hinge & beta1-3 plus major TURN

|               |    |             |                      |                |                                     |                                |                               |
|---------------|----|-------------|----------------------|----------------|-------------------------------------|--------------------------------|-------------------------------|
| Synechocystis | 6  | LKFRFRFIMTQ | ISGSPDVPDLGRQFMNLL   | TFGTITGVAAGALY | PAVKYLIPPPSSGGGGVTAKDALGNDVKV       | TEFLASHNA                      | GDRVLAQGLKQDP                 |
| Nitrospira    | 18 | VGSRRTPFHW  | VTVAATAAVVGVGLAVPLLG | SLISPAFMRRR    | RAWVDVGPVDELPAQRPTQLDHTVTTVRDGN     | QEAATSHKAV                     |                               |
| Epsilon       | 3  | DIQRDFPLGM  | SLASVTAIGAIASLVAMKK  | TWDPSPSVVSAG   | FTTIDVANMQEGQFSTVEWRCKPVYILKRSK     | KEGFNNEK                       | RDPKFI                        |
| Kuenenia_2    | 21 | TISRRNFLT   | AGWGLFFAAIGAYLSQLFG  | YKGFYPKVLF     | EPSRPFVVEGPKQFPVGSVTTLKSRIKIFVVRDAD |                                |                               |
| Schlesneria_2 | 12 | TPPRRTFLAA  | ISSVVIGGFITLTPLVSLG  | LAFFLDPLVLRTR  | TKFKGGDSEGFLLFVANLSDLPDDGTPLRFL     | DDGTPLRFL                      | OMRARDKQDAWNVV                |
| Methylocy_2   | 8  | DPLRRDVLVY  | ATGACAVVATAAAIWPFLD  | ALNPDAS        | SVAAGA                              | PIDVDLGGLEPGKVVVRWRLPVAIFRRP   | DAALEKLRKDPVLL                |
| Saccharib     | 22 | RPTRRTVLTA  | GSVALGGAGACTLAAPFLD  | SLRRTRAS       | AASTQEET                            | VLDISVQDLEPGAHKIVVWQMPVAIQRRTP | EMIRALEDPALLA                 |
| Glu_oxyd      | 20 | GSRRRDVLAT  | VTVAMGCAGACAVVYFPLD  | SLNGTRAN       | HVGESD                              | ILDVDLSTLKAGQQIVVTVWRGVPVFORRT | PEMLKTLQDPAILLO               |
| Gluc_dia      | 20 | GTGRRDFLTL  | VTLAAGGAGAAAFANWFLD  | SLRPGDTR       | SARA                                | PVDVDVSKLPPGQQITVVMHGSVPVFTI   | TRTPESLARLQAAALAA             |
| Nitrosomonas  | 10 | MSGRRRFLV   | ATSVAGAVAGAVATPFLR   | SMP            | SERA                                | KAAGA                          | PVEVDISKLEPGMLLTAEWRGKVVVWLRK |
| Beta          | 6  | NNGRRRFLTL  | ATCGAGGVAALGVATPFAV  | SFFP           | SEKA                                | K-AGA                          | AVEVDVSKIEAGQLLTAEWQCKPIWVLR  |
| Nitrosoc      | 7  | DQKRRRFLTA  | ATTVVGVMGFTAVVFFVK   | SMQP           | SERA                                | QAAGA                          | PVEVDISNLEPGHIMTVWRGKPVWILRR  |
| Frateuria     | 6  | DGRRRFLAV   | STVSVGGAGLAAAAPFAE   | SMQP           | SARA                                | RSLAA                          | PVQVDLAGIAQEQMVTYAWRGQPVFV    |
| Nitrobacter   | 17 | SPTRRDFLYI  | ATAAVGTGVAIAIWPMSI   | QLEP           | DKAT                                | LAAGG                          | PVELDLSPVQPGQVSVRWRDHPILVINR  |
| Methylob_r    | 13 | DGSRRDPLFL  | ATGAGLAVGAGAAVWPLIS  | SMAP           | DADT                                | IAAGA                          | PTEVDLTPIQDGGQIVNVFWRGKLI     |
| Rickettsia    | 11 | QTRRRDFMVL  | TASSVAAVGAVCTLWPLVD  | SLNP           | SADV                                | LALS                           | SIEVDLSNIAVGGQIVTVWRGKPVF     |
| Maricaulis_1  | 14 | EPSRRDFLYI  | AAGGVAAVGGGLVANWFLD  | QMPN           | AADT                                | LALG                           | SIRVDVSVVPGSEIITVMWRGSPV      |
| RoseobacterA2 | 8  | EGTRRDFLYY  | ATAGAGAVTAGAAIWPVNV  | QMPN           | TADV                                | LALS                           | SIRVDVSGIEPGTQLTVKWLKPV       |
| Magnetococcus | 8  | DENRRDPLIM  | ATGAVGAVGAVAGAAWPLD  | SWSP           | SADV                                | LALS                           | TTEVSVAAIAEGQMITMVGKPVF       |
| Methylocy_1   | 12 | EPSRRDILYI  | ATGAAAGAAAGMVTWPLIS  | SMNT           | DAST                                | LALA                           | STEVDLSSVEEGQIVTVWRGKPV       |
| Beijerinckia  | 8  | APTTRDFLFI  | ATGAMGVVAAGSVLWPLD   | QMPN           | DAST                                | LAAA                           | TTEVDIGNIAEGQIVTVWRGKPV       |
| Micavibrio    | 16 | GTTRRDFLFL  | TAGAFGAVGAGIFSFWPLD  | SMNP           | AADT                                | LAMA                           | STEVDLSPVEVGSQITVTVWRG        |

|              |    |                        |            |      |      |      |                         |
|--------------|----|------------------------|------------|------|------|------|-------------------------|
| Acanthamoeba | 62 | ESGGRAFSYMLVGTAGVGYAAA | AKHTVTKFLD | SMNP | AANV | RAMA | NVEVDVSNIAEGTITMIVKWR   |
| S_cerevisiae | 50 | ADKGRSYAYFMVGMALSSAGAK | STVETFIAS  | SMNT | DAST | LALA | KVEVNLAAIPLCKNVVVKQCKPV |
| Chicken      | 29 | SEDRKGFSYLVTATACVATAYA | AKNVVTFQIS | SLSA | SADV | LALS | KIEIKLSDIPEGKNVAFKWR    |

Features RR 1  $\alpha$  turn less in bacteria Cimit1 Cimit2

Rieske ISP, C-terminal part: Iron Sulfur cluster domain to C terminus

|               |            |            |                  |        |           |            |              |           |        |        |         |         |        |        |
|---------------|------------|------------|------------------|--------|-----------|------------|--------------|-----------|--------|--------|---------|---------|--------|--------|
| Synechocystis |            | TYIVVQDDTI | IANYGINAVCTHLGCV | PW     | NAS       | ENKFMCPCHG | SOYNAEGKVVRC | PAPLSLALA | HATVTD | DDKLVL | STWTET  | DFRDEDP | WVA    | 192    |
| Nitrospira    |            | WAVKQPGDVT | VPSPICTHLGC      | YRW    | DDA       | EKKFLCPC   | HSSFDVKGDV   | LG        | PAPRPL | DLRPAK | VEGGRLL | VMYKDFK | SGLRDS | VEL    |
| Epsilon       | 79         | GGSVPTAA   | IQICTHLGCI       | PTIY   | QDE       | EKFLCPC    | HGRFSD       | GNI       | AGT    | PPRPF  | DI      | PPFKI   | EGTKIT | FG     |
| Kuenenia_2    |            | VRDADS     | FKAISVVCTHLG     | CAVEF  | SKE       | NIFCEPC    | HGSKYR       | NGVNFAG   | PAPRPL | DHLQ   | MIL     | DNNGK   | LAVDT  | SVKVP  |
| Schlesneria_2 | ANQT       | IGTIYLR    | KMPGNQII         | AFNDT  | CPHLG     | CKVDY      | QENS         | KSLFCPC   | HASAFK | LDGERT | NKI     | PPRGL   | DLT    | SKTDT  |
| Methylocy_2   | DSEVL-QQP  | RYARNW     | HRSIDPT          | YAVLVG | VCTHLG    | CIPLYP     | TPSGA        | EPTAD     | WP     | GGFFC  | PHGSK   | FDL     | AGRVYK | GVP    |
| Saccharib     | QSRIR-QQP  | REAENA     | HRSLR            | PDYGVV | TICTHLG   | CVDPYQ     | NITS         | LPAS      | GGFFC  | PHGSK  | FDG     | AGRVL   | RNM    | PAPYNL |
| Glu_oxyd      | ESCIF-QQP  | KDATN      | WRSVSP           | DIGVM  | IGICTHLG  | CVPTFD     | APTQ         | AEP       | ACKYLC | PHGSK  | FD      | S       | AGRAYR | NAP    |
| Gluc_dia      | SSGAL-QQP  | PYAAN      | HRSIK            | PEFV   | VVGICTHLG | CVPTYS     | PADP         | ATIAN     | WP     | GGYAC  | PHGSK   | FD      | LAGRV  | FT     |
| Nitrosomonas  | DSQRD-QQP  | PYAQN      | HRSIK            | PELL   | VLVG      | VCTHLG     | CSVYR        | KDI       | APADL  | GS     | SDWL    | GGFFC   | PHGSK  | FD     |
| Beta          | NSDAE-QQP  | EYAKN      | HRSIK            | PELL   | VVIG      | ICTHLG     | CSPTFR       | PDIA      | PADL   | GD     | ADWK    | GGFFC   | PHGSK  | FD     |
| Nitrosoc      | QSEVES-QQP | FAENI      | YRSIK            | PELL   | VVIG      | ICTHLG     | CSPSFR       | PEV       | EDL    | GD     | PDW     | GGFFC   | PHGSK  | FD     |
| Frateuria     | DSKVD-QQP  | AWARN      | HRSQR            | PAWL   | VLVGL     | CTHLG      | CVDPYV       | YV        | GS     | LG     | RE      | PF      | PDW    | GGFFC  |
| Nitrobacter   | QSSEF-QQP  | EYAHN      | HRSAN            | PEY    | GVF       | VIGICTHLG  | CIPLYP       | SP        | NAN    | H      | PA      | SD      | W      | GGYFC  |
| Methylob_r    | AAFFT      | VKS        | GH               | Q      | WLVV      | YGNCTHLG   | CVPIGH       | QGN       | F      | EG     | W       | AC      | PHGSK  | FD     |
| Rickettsia    | ETDEA      | RVK        | AGH              | DM     | LV        | TI         | GICTHLG      | CVPLAN    | Q      | E      | Y       | D       | G      | W      |
| RoseobacterA2 | LAENANI    | APDA       | DA               | SD     | ENR       | SL         | DES          | GE        | LV     | MM     | GV      | CTH     | FC     | VP     |
| Maricaulis_1  | DARND-NL   | PAGV       | CP               | TD     | ANR       | VLR        | PE           | F         | L      | V      | K       | GN      | CTH    | GC     |
| Magnetococcus | QADAD      | R          | V                | K      | DE        | W          | L            | V         | L      | A      | I       | C       | T      | H      |
| Micavibrio    | QTD        | E          | D                | R      | A         | T          | K            | PE        | W      | L      | V       | I       | V      | I      |
| Methylocy_1   | E          | P          | D                | A      | K         | R          | V            | O         | K      | PE     | W       | L       | V      | I      |
| Beijerinckia  | Q          | D          | O                | S      | R         | V          | O            | K         | PE     | W      | L       | V       | I      | V      |

|              |     |        |        |     |   |    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |   |   |         |   |     |
|--------------|-----|--------|--------|-----|---|----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|-----|---|---|---------|---|-----|
| Acanthamoeba |     | CPDAAN | RRKADK | PEW | L | V  | L | V | G | V | C | T | H | G | C | V | P | L | G | Q | E | Y | G | G | W | F | C | P | H | G | S | H | Y | D | T | A | G | R | I | R | G | P | A | P | L | N | L | E | V | P | P | Y | V | F | K | D | S | K | V | L | V | G   | V | D | S       | V | 235 |
| S_cerevisiae |     | QTDADR | V      | K   | D | PE | W | L | I | L | G | I | C | T | H | G | C | V | P | L | G | A | G | D | F | G | W | F | C | P | H | G | S | H | Y | D | T | A | G | R | I | R | G | P | A | P | L | N | L | E | P | A | Y | E | F | D | G | K | V | I | V | 185 |   |   |         |   |     |
| Chicken      | 121 | QHLD   | R      | V   | K | PE | W | L | V | L | V | G | I | C | T | H | G | C | V | P | L | G | A | S | G | D | F | G | Y | Y | F | C | P | H | G | S | H | Y | D | T | A | G | R | I | R | G | P | A | P | N | L | E | V | P | T | Y | Q | V | G | D | L | V   | V | G | 196/272 |   |     |

features Bac Insert +H+G+P  $\alpha$ Insert G C H S Y165 C-terminus  
 Con.Indels Cimit 3a 3b Cimit4 Cimit5 Cimit6 Cimit7



## Supporting Information Tables

These tables contain large datasets that do not fit the recommended format of the text and are therefore presented as pasted images; they are also available as Excel files, if requested.

### **Table S1 – Genomic distribution of bioenergetic systems in $\alpha$ -proteobacteria.**

**A.** The genomes of *ca.* 120  $\alpha$ -proteobacterial organisms were studied from the latest version of the genome NCBI database <http://www.ncbi.nlm.nih.gov/genome/browse/> - accessed on 14 March 2014, verifying also the completeness of genomic data (\*). Reconstruction of the various bioenergetic systems (see text) was deduced by combining genomic information with biochemical and microbiological data. The organisms are listed following the left-right sequence in the model of Fig. 1B. Two major types of *bd* oxidases are classified as *bd*-I or CIO [25,79]. The organisms directly shown in Fig. 1B are **yellow highlighted** and those proposed to be relatives of mitochondria are shown in italics with pertinent references (including [86,87]). Underlined organisms are symbionts or pathogens. **B.** This table lists the organisms that have been analysed but are not included in the model of Fig. 1B, also because they are in parallel paths of evolution with respect to the mitochondrial subset of bioenergetic systems. The organisms highlighted in pale yellow are shown in Fig. 1, while other annotations are the same as in **A**. Complementary information is in Table S2.

A

| Table S1a | Organism/Name - underlined, endocellular/parasite<br>proposed mitochondrial relative with Reference<br>Shown in Fig. 1B | Order/Group         | BIONERGETIC SYSTEMS |        |               |         |       | NOTES   | Genome*<br>complete | Photosynthetic |
|-----------|---|---------------------|---------------------|--------|---------------|---------|-------|---|---------------------|----------------|
|           |   |                     | 1. bc1              | 4. aa3 | 2. bd -types  | 5. cbb3 | 3. bo |   |                     |                |
|           | All 6 bioenergetic systems  |                     |                     |        |               |         |       |   |                     |                |
|           | <i>Paracoccus denitrificans</i> PD1223 [17]   | Rhodobacteriales    | YES                 | YES    | YES           | YES     | YES   | Nitrate and nitrite reductase AND NOR; CIO bd oxidase; <i>ba3</i> as <i>bo</i> oxidase                          | YES                 |                |
|           | <i>Aerobacterium tumefaciens</i> E2 [86]  | Rhizobiales         | YES                 | YES    | YES           | YES     | YES   | Nitrate and Nitrite reductase, NOR  | YES                 |                |
|           | <i>Ochrobactrum anthropi</i> ATCC 49188 [6,7]   | Rhizobiales         | YES                 | YES    | YES           | YES     | YES   | Nitrate and Nitrite reductase, NOR  | YES                 |                |
|           | <i>Bradyrhizobium</i> ORS J75 [20]  | Rhizobiales         | YES                 | YES    | YES           | YES     | YES   | Nitrate and Nitrite reductase, NOR like oxidase; CIO bd oxidase; <b>2 ISP forms</b>                             | YES                 | YES            |
|           | <i>Rhodospseudomonas palustris</i> HoA2 [6,7]   | Rhizobiales         | YES                 | YES    | YES - 2 types | YES     | YES   | Nitrate and Nitrite reductase, NOR; CIO bd oxidase; <b>2 ISP forms</b>  | YES                 | YES            |
|           | <i>Methylobacterium nodulans</i>  | Rhizobiales         | YES                 | YES    | YES           | YES     | YES   | Nitrate and Nitrite reductase; CIO bd oxidase   | YES                 |                |
|           | <i>Microvirga</i> sp. WSM3557   | Rhizobiales         | YES                 | YES    | YES - 2 types | YES     | YES   | Multiple COX operons; Nitrate and Nitrite reductase; CIO bd oxidase   | YES                 |                |
|           | <i>Rhizobium leguminosarum</i> bv. <i>trifolii</i> WSM2304  | Rhizobiales         | YES                 | YES    | YES           | YES     | YES   | Nitrate and nitrite reductase AND NOR   | YES                 |                |
|           | <i>Mesorhizobium loti</i> MAFF303099  | Rhizobiales         | YES                 | YES    | YES           | YES     | YES   | Nitrate and Nitrite reductase, NOR; CIO bd oxidase  | YES                 |                |
|           | <i>Brucella suis</i> 1330   | Rhizobiales         | YES                 | YES    | YES           | YES     | YES   | Nitrate and nitrite reductase AND NOR   | YES                 |                |
|           | <i>Methylocella silvestris</i>  | Rhizobiales         | YES                 | YES    | YES           | YES     | YES   | Methane monooxygenase; Nitrate reductase, <b>assimilatory Nitrite reductase</b> ; CIO bd oxidase                | YES                 | YES            |
|           | <i>Hyphomicrobium denitrificans</i> ATCC 51888  | Rhizobiales         | YES                 | YES    | YES           | YES     | YES   | Nitrate and Nitrite reductase AND NOR; CIO bd oxidase   | YES                 | YES            |
|           | <i>Parvibaculum lavamentivorans</i> DS-1  | Rhizobiales         | YES                 | YES    | YES           | YES     | YES   | Nitrate and nitrite reductase AND NOR; CIO bd oxidase   | YES                 |                |
|           | <i>Tistrella mobilis</i> KA081020-065 [48]  | Rhodospirillales    | YES                 | YES    | YES - 2 types | YES     | YES   | Nitrate and Nitrite reductase; CIO bd oxidase   | YES                 |                |
|           | <i>Thalassospira profundimaris</i> WP0211   | Rhodospirillales    | YES                 | YES    | YES           | YES     | YES   | Nitrate and Nitrite reductase; CIO bd oxidase   | NO                  |                |
|           | <i>Rhodospirillum centenum</i> SW   | Rhodospirillales    | YES                 | YES    | YES - 3 types | YES     | YES   | Nitrate reductase periplasmic and Nitrite reductase; CIO bd oxidase; <b>2 ISP forms</b>                         | YES                 | YES            |
|           | <i>Caulobacter</i> [18]   | Caulobacteriales    | YES                 | YES    | YES           | YES     | YES   | Nitrate and Nitrite reductase   | YES                 |                |
|           | <i>Asticcacaulis excentricus</i>  | Caulobacteriales    | YES                 | YES    | YES           | YES     | YES   | Nitrate and Nitrite reductase AND NOR   | YES                 |                |
|           | <i>Novosphingobium pentaromatorans</i><br>minus 1, bd   | Sphingomonadales    | YES                 | YES    | YES           | YES     | YES   | Nitrate and Nitrite reductase   | YES                 |                |
|           | <i>Sphingopyxis alaskensis</i> RB2256   | Sphingomonadales    | YES                 | YES    | NO            | YES     | YES   | <b>assimilatory Nitrate &amp; Nitrite reductase</b>   | YES                 |                |
|           | <i>Gluconacetobacter europaeus</i> LMG 18494  | Rhodospirillales    | YES                 | YES    | YES - 3 types | NO      | YES   | Nitrate and Nitrite reductase; <b>2</b> CIO bd oxidase; <b>2 ISP forms</b>                                      | YES                 |                |
|           | <i>Gluconacetobacter diazotrophicus</i> PAI 5   | Rhodospirillales    | YES                 | YES    | YES           | NO      | YES   | Nitrate and Nitrite reductase; CIO bd oxidase; <b>2 ISP forms</b>   | YES                 |                |
|           | <i>Azospirillum lipoferum</i> 4B  | Rhodospirillales    | YES                 | YES    | YES           | NO      | YES   | COX not validated as aa3; Nitrite reductase; CIO bd oxidase   | YES                 |                |
|           | <i>Acetobacteraceae</i> bacterium AT-5844   | Rhodospirillales    | YES                 | YES    | YES - 2 types | NO      | YES   | Nitrate and Nitrite reductase; CIO bd oxidase   | NO                  |                |
|           | <i>Methylocystis</i> SC2  | Rhizobiales         | YES                 | YES    | YES - 2 bd-I  | NO      | YES   | Multiple COX operons; Methane monooxygenases; Nitrate and Nitrite reductase; <b>2 ISP forms</b>                 | YES                 |                |
|           | <i>Methylobacterium radiotolerans</i> JCM 2831  | Rhizobiales         | YES                 | YES    | YES           | NO      | YES   | Nitrate reductase periplasmic; Nitrite reductase; <b>3</b> CIO bd oxidases                                      | YES                 | YES            |
|           | <i>Methylobacterium extorquens</i> (4 strains)  | Rhizobiales         | YES                 | YES    | YES - 2 types | NO      | YES   | Nitrate reductase periplasmic; Nitrite reductase; <b>2</b> or <b>3</b> CIO bd oxidase                           | YES                 | YES            |
|           | <i>Methylosinus trichosporium</i>   | Rhizobiales         | YES                 | YES    | YES           | NO      | YES   | Methane monooxygenase; Nitrate and Nitrite reductase  | NO                  |                |
|           | <i>Nitrobacter hamburgensis</i> X14   | Rhizobiales         | YES                 | YES    | YES           | NO      | YES   | Nitrite oxidoreductase and <b>assimilatory reductase</b> , NOR like oxidase; CIO bd oxidase; <b>2 ISP forms</b> | YES                 | YES            |
|           | <i>Azorhizobium caulinodans</i> ORS 571   | Rhizobiales         | YES                 | YES    | YES - 2 types | NO      | YES   | Nitrite reductase; CIO bd oxidase   | YES                 |                |
|           | <i>Sinorhizobium meliloti</i> AK83  | Rhizobiales         | YES                 | YES    | YES           | NO      | YES   | COX not validated as aa3; Nitrate reductase periplasmic, Nitrite reductase                                      | YES                 |                |
|           | <i>Phyllobacterium</i> sp. YR531  | Rhizobiales         | YES                 | YES    | YES           | NO      | YES   | Nitrate and Nitrite reductase   | YES                 |                |
|           | <i>Pelagibacterium halotolerans</i> B2  | Rhizobiales         | YES                 | YES    | YES           | NO      | YES   | Nitrate and Nitrite reductase   | YES                 |                |
|           | <i>Parvularcula bermudensis</i> HTCC2503 [18]   | Parvularculales     | YES                 | YES    | YES           | NO      | YES   | Nitrate reductase; CIO bd oxidase; also subunit II of a non functional CIO bd oxidase                           | YES                 |                |
|           | <i>Rhodobacter sphaeroides</i> ATCC 17025 [6-8]   | Rhodobacteriales    | YES                 | YES    | YES           | YES     | NO    | Nitrate and Nitrite reductase AND NOR; CIO bd oxidase; <i>bo</i> present in other strains                       | YES                 | YES            |
|           | <i>Rhodospseudomonas palustris</i> BisA53 [6,7]   | Rhizobiales         | YES                 | YES    | YES           | YES     | NO    | Nitrate and Nitrite reductase, NOR; <b>3 ISP forms</b> ; CIO bd oxidase   | YES                 | YES            |
|           | <i>Xanthobacter autotrophicus</i>   | Rhizobiales         | YES                 | YES    | YES           | YES     | NO    | Nitrate reductase and <b>assimilatory Nitrite reductase</b> ; CIO bd oxidase                                    | YES                 |                |
|           | <i>Afipia felis</i>   | Rhizobiales         | YES                 | YES    | YES           | YES     | NO    | Nitrate and Nitrite reductase, NOR like <i>cyt c</i> oxidase; CIO bd oxidase; <b>2 ISP forms</b>                | NO                  |                |
|           | <i>Brucella abortus</i> S19   | Rhizobiales         | YES                 | YES    | YES           | YES     | NO    | Nitrite reductase only  | YES                 |                |
|           | <i>Caenispirillum salinarum</i> AK4   | Rhodospirillales    | YES                 | YES    | YES           | YES     | NO    | Nitrate and Nitrite reductase, NOR; <b>2</b> CIO bd oxidases; <b>2 ISP forms</b>                                | NO                  | YES            |
|           | <i>Sphingobium</i> sp. SYK-6  | Sphingomonadales    | YES                 | YES    | YES           | YES     | NO    | <i>bo</i> oxidase present in other <i>Sphingobium</i> species; assimilatory Nitrate and Nitrite reductase       | YES                 | YES            |
|           | <i>Maricaulis maris</i>   | Rhodobacteriales    | YES                 | YES    | NO            | YES     | NO    | Nitrate reductase, respiratory; <b>2 ISP forms</b>  | YES                 |                |
|           | <i>Roseobacter litoralis</i>  | Rhodobacteriales    | YES                 | YES    | NO            | YES     | NO    | Nitrate reductase and assimilatory Nitrite reductase; <b>3 ISP forms</b>  | YES                 | YES            |
|           | <i>Magnetospirillum magneticum</i> AMB-1  | Rhodospirillales    | YES                 | YES    | NO            | YES     | NO    | Nitrate reductase periplasmic, <b>assimilatory Nitrite reductase</b> ; NOR                                      | YES                 |                |
|           | <i>Micavibrio aeruginosavorus</i> ARL-13 [19]   | Rhodospirillales?   | YES                 | YES    | YES           | NO      | NO    | Ectoparasite of bacteria; Nitrate reductase, respiratory; CIO bd oxidase  | YES                 |                |
|           | <i>Beijerinckia indica</i> subsp. <i>indica</i>   | Rhizobiales         | YES                 | YES    | YES           | NO      | NO    | Nitrate reductase and <b>assimilatory Nitrite reductase</b> ; CIO bd oxidase                                    | YES                 |                |
|           | <i>Rhodospirillum rubrum</i> [20]   | Rhodospirillales    | YES                 | NO     | YES           | YES     | NO    | Nitrite reductase in various strains of the species; NOR; <b>2 ISP forms</b>                                    | YES                 | YES            |
|           | <i>Rhodospirillum photometricum</i> DSM 122   | Rhodospirillales    | YES                 | NO     | YES           | YES     | NO    | Nitrite reductase only; <b>2 ISP forms</b>  | YES                 | YES            |
|           | <i>Oligotropha carboxidovorans</i> OM4  | Rhizobiales         | YES                 | NO     | YES           | YES     | NO    | Nitrate reductase respiratory, Nitrite reductase; CIO bd oxidase  | YES                 |                |
|           | <i>Rickettsia prowazekii</i> Bp22 [4-7,87]  | Rickettsiales       | YES                 | YES    | YES           | NO      | NO    | <b>NO N</b> metabolism; CIO bd oxidase  | YES                 |                |
|           | <i>Rickettsia felis</i> URRW/Col2 [4-7,88]  | Rickettsiales       | YES                 | YES    | YES           | NO      | NO    | <b>NO N</b> metabolism; CIO bd oxidase  | YES                 |                |
|           | <i>Wolbachia</i> sp. wRI  | Rickettsiales       | YES                 | YES    | YES           | NO      | NO    | <b>NO N</b> metabolism; CIO bd oxidase  | YES                 |                |
|           | <i>Wolbachia</i> endosymbiont of <i>Drosophila melanogaster</i>   | Rickettsiales       | YES                 | YES    | YES           | NO      | NO    | <b>NO N</b> metabolism; CIO bd oxidase  | YES                 |                |
|           | <i>Candidatus Odysseella thessalonicensis</i> L13   | Rickettsiales       | YES                 | YES    | YES           | NO      | NO    | <b>NO N</b> metabolism  | NO                  |                |
|           | <i>alpha</i> proteobacterium BAL199   | Rickettsiales       | YES                 | YES    | YES           | NO      | NO    | <b>NO N</b> ; CIO bd oxidase  | NO                  |                |
|           | Mitochondrion, fungi & protists   | eukaryotes          | YES                 | YES    | NO            | NO      | NO    | <b>Assimilatory Nitrate and Nitrite reductases</b>  | YES                 |                |
|           | <i>Magnetococcus marinus</i> MC-1   | Magnetococcales     | YES                 | NO     | NO            | YES     | NO    | COX1 similar to NOR, but COX2 and COX3 missing; Nitrate reductase periplasmic                                   | YES                 |                |
|           | <i>Candidatus Mithochondrii iric VA</i> [21]  | Rickettsiales       | YES                 | NO     | YES           | YES     | NO    | <b>NO N</b> metabolism; CIO bd oxidase  | YES                 |                |
|           | <i>Candidatus Endolissoclinum faulkneri</i> LS  | Rhodospirillales    | YES                 | NO     | NO            | YES     | NO    | Nitrite reductase   | YES                 |                |
|           | <i>Anaplasma marginale</i> str. <i>St. Maries</i>   | Rickettsiales Anap. | YES                 | YES    | NO            | NO      | NO    | <b>NO N</b> metabolism  | YES                 |                |
|           | <i>Wolbachia</i> endosymbiont of <i>Culex quinquefasciatus</i> NIV  | Rickettsiales Anap. | YES                 | YES    | NO            | NO      | NO    | <b>NO N</b> metabolism  | YES                 |                |
|           | <i>Ehrlichia canis</i> str. <i>Jake</i>   | Rickettsiales Anap. | YES                 | YES    | NO            | NO      | NO    | <b>NO N</b> metabolism  | YES                 |                |
|           | <i>Neorickettsia risticii</i> str. <i>Illinois</i>  | Rickettsiales Anap. | YES                 | YES    | NO            | NO      | NO    | <b>NO N</b> metabolism  | YES                 |                |
|           | <i>Candidatus Pelagibacter ubique</i> HTCC1062 [13,14]  | SAR1 clade          | YES                 | YES    | NO            | NO      | NO    | <b>NO N</b> metabolism  | YES                 |                |
|           | <i>Candidatus Puniceispirillum marinum</i> IMCC1322 [13,14]   | SAR11 clade         | YES                 | YES    | NO            | NO      | NO    | <b>NO N</b> metabolism - a sulfite reductase remotely related to assimilatory nitrite reductases                | YES                 |                |
|           | <i>SAR116 cluster</i> HIMB100 [13]  | SAR11 clade         | YES                 | YES    | NO            | NO      | NO    | <b>NO N</b> metabolism  | NO                  |                |
|           | Mitochondrion - metazoans   | eukaryotes          | YES                 | YES    | NO            | NO      | NO    | <b>NO N</b> metabolism  | YES                 |                |

**B**

| Table S1b   |                  | BIOENERGETIC SYSTEMS |        |               |         |       | Nitrogen (N) metabolism [see also Table 1]; NOR, nitric oxide reductase; CIO, Cyanide-insensitive bd oxidase [79]   |  | Genome* complete | Photosynthetic |
|---|------------------|----------------------|--------|---------------|---------|-------|---|--|------------------|----------------|
| Organism/Name shown in Fig. S1                              | Order/Group      | 1. bc1               | 4. aa3 | 2. bd -types  | 5. cbb3 | 3. bo |   |  |                  |                |
| <b>All 6 bioenergetic systems</b>                           |                  |                      |        |               |         |       |   |  |                  |                |
| <i>Starkeya novella</i> DSM 506                             | Rhizobiales      | YES                  | YES    | YES           | YES     | YES   | Nitrate and Nitrite reductase; CIO bd oxidase   |  | YES              |                |
| <i>Rhodovulum</i> sp. PH10                                  | Rhodobacteriales | YES                  | YES    | YES - 2 types | YES     | YES   | Nitrite reductase; 2 CIO bd oxidases; <b>3 ISP forms</b>  |  | NO               |                |
| <i>Azospirillum</i> sp. B510                                | Rhodospirillales | YES                  | YES    | YES - bd-I    | YES     | YES   | Nitrate reductase periplasmic and Nitrite reductase   |  | YES              |                |
| <i>Roseomonas cervicalis</i> ATCC 49957                     | Rhodospirillales | YES                  | YES    | YES - 3 types | YES     | YES   | NOR; 2 CIO bd oxidases  |  | NO               |                |
| <i>Thalassospira xiamenensis</i> M-5 = DSM 17429            | Rhodospirillales | YES                  | YES    | YES           | YES     | YES   | Nitrate reductase periplasmic, <b>assimilatory Nitrite reductase</b> ; CIO bd oxidase                               |  | NO               |                |
| <b>minus 1</b>  |                  |                      |        |               |         |       |   |  |                  |                |
| <i>Azospirillum amazonense</i> Y2                           | Rhodospirillales | NO                   | YES    | YES - 2 types | YES     | YES   | No ISP in bc1; Nitrate reductase periplasmic, Nitrite reductase; CIO bd oxidase                                     |  | NO               |                |
| <i>Azospirillum brasilense</i> Sp245                        | Rhodospirillales | NO                   | YES    | YES           | YES     | YES   | NO cytochrome b nor c1; Nitrate reductase periplasmic, Nitrite reductase; CIO bd oxidase                            |  | YES              |                |
| <i>Roseomonas</i> sp. B5                                    | Rhodospirillales | YES                  | YES    | YES - 2 types | NO      | YES   | Nitrate and Nitrite reductase; CIO bd oxidase   |  | NO               |                |
| <i>Magnetospirillum gryphiswaldense</i> MSR-1 v2            | Rhodospirillales | YES                  | YES    | YES           | YES     | NO?   | COX possibly working also as ubiquinol oxidase; Nitrate and nitrite reductase; CIO bd oxidase                       |  | YES              |                |
| <i>Oceanibaculum indicum</i> P24                            | Rhodospirillales | YES                  | YES    | YES - bd-I    | YES     | NO    | Nitrate reductase   |  | NO               |                |
| <i>Gluconacetobacter oboediensis</i> 1748p2                 | Rhodospirillales | YES                  | NO     | YES - 3 types | YES     | YES   | <b>assimilatory Nitrite reductase</b> ; 2 CIO bd oxidases; <b>2 ISP forms</b>                                       |  | NO               |                |
| <i>Fulvimarina pelagi</i>                                   | Rhizobiales      | YES                  | YES    | YES           | NO      | YES   | Nitrate and Nitrite reductase; CIO bd oxidase   |  | NO               |                |
| <i>Polymorphum gilvum</i> SL003B-26A1                       | Polymorphum      | YES                  | YES    | YES - 2 types | YES     | NO    | Nitrate and Nitrite reductase; CIO bd oxidase   |  | YES              |                |
| <b>minus 2</b>  |                  |                      |        |               |         |       |   |  |                  |                |
| <i>Acetobacter pomorum</i>                                  | Rhodospirillales | YES                  | NO     | YES - 2 types | NO      | YES   | Nitrate reductase; CIO bd oxidase   |  | NO               |                |
| <i>Acetobacter pasteurianus</i> IFO 3283-01-42C (9 strains) | Rhodospirillales | YES                  | NO     | YES - 2 types | NO      | YES   | Nitrate reductase and <b>assimilatory Nitrite reductase</b> ; CIO bd oxidase  |  | YES              |                |
| <i>Gluconacetobacter xylinus</i> NBRC 3288                  | Rhodospirillales | YES                  | NO     | YES           | NO      | YES   | Nitrate and Nitrite reductase; CIO bd oxidase; <b>2 ISP forms</b>   |  | YES              |                |
| <i>Gluconobacter oxydans</i> 621H                           | Rhodospirillales | YES                  | NO     | YES           | NO      | YES   | COX not functional; <b>assimilatory Nitrite reductase</b> ; CIO bd oxidase  |  | YES              |                |
| <i>Gluconobacter morbifer</i>                               | Rhodospirillales | YES                  | NO     | YES           | NO      | YES   | COX not functional; Nitrite reductase; 2 CIO bd oxidases  |  | NO               |                |
| <i>Saccharibacter</i> sp.                                   | Rhodospirillales | YES                  | NO     | YES           | NO      | YES   | COX not functional; Nitrite reductase; CIO bd oxidase   |  | YES              |                |
| <i>Saccharibacter floricola</i>                             | Rhodospirillales | YES                  | NO     | YES - bd-I    | NO      | YES   | COX not functional; Nitrite reductase   |  | NO               |                |
| <i>Commensalibacter intestini</i> A911                      | Rhodospirillales | YES                  | NO     | YES - bd-I    | NO      | YES   | COX not functional; Nitrite reductase; CIO bd oxidase; <b>2 ISP forms</b>   |  | NO               |                |
| <i>Acidiphilium cryptum</i> JF-5                            | Rhodospirillales | NO                   | YES    | YES 2 types   | NO      | YES   | cytochrome b deranged and NO ISP; <b>assimilatory Nitrate and Nitrite reductase</b> ; CIO bd oxidase                |  | YES              |                |
| <i>Acidocella</i>   | Rhodospirillales | NO                   | YES    | YES 2 types   | NO      | YES   | cytochrome b & ISP deranged; respiratory Nitrate reductase & <b>assimilatory Nitrite reductase</b> ; CIO bd oxidase |  | YES              |                |
| <i>Granulibacter thebesdensis</i> CGDNIH1                   | Rhodospirillales | NO                   | YES    | YES           | NO      | YES   | <b>assimilatory Nitrate and Nitrite reductase</b> ; CIO bd oxidase  |  | YES              |                |
| <i>Brevundimonas diminita</i>                               | Caulobacteriales | YES                  | YES    | YES - bd-I    | NO      | YES   | NO N metabolism   |  | NO               |                |
| <i>Phaeospirillum molischianum</i>                          | Rhodospirillales | YES                  | NO     | YES - bd-I    | YES     | NO    | Nitrite reductase only; <b>2 ISP forms</b>  |  | YES              |                |
| <i>Nitratireductor indicus</i> C115                         | Rhizobiales      | YES                  | NO     | YES - bd-I    | YES     | NO    | <b>assimilatory Nitrate reductase</b> , Nitrite reductase   |  | NO               |                |
| <i>Rhodomicrobium vannielii</i> ATCC 17100                  | Rhizobiales      | YES                  | NO     | YES           | YES     | NO    | Nitrite reductase; CIO bd oxidase   |  | YES              |                |
| <i>Magnetospirillum magnetotacticum</i>                     | Rhodospirillales | YES                  | YES    | NO            | YES     | NO?   | COX possibly working also as ubiquinol oxidase; Nitrate and nitrite reductase                                       |  | NO               |                |
| <b>minus 3</b>  |                  |                      |        |               |         |       |   |  |                  |                |
| <i>Acetobacter aceti</i>                                    | Rhodospirillales | NO                   | NO     | YES           | NO      | YES   | COX 1 short; Nitrite reductase; 2 CIO bd oxidases, 1 partial  |  | NO               |                |
| <i>Asaia platycody</i>                                      | Rhodospirillales | NO                   | NO     | YES - bd-I    | NO      | YES   | COX not functional; <b>assimilatory Nitrite reductase</b>   |  | NO               |                |
| <i>Asaia</i> sp. SFZ.1                                      | Rhodospirillales | NO                   | NO     | YES           | NO      | YES   | nitrite and nitrate reductase; CIO bd oxidase   |  | NO               |                |
| <i>Gluconacetobacter hansenii</i> ATCC 23769                | Rhodospirillales | YES                  | NO     | NO            | NO      | YES   | Nitrite reductase   |  | YES              |                |
| <i>Zymomonas mobilis</i> subsp. <i>mobilis</i> ZM4          | Sphingomonadales | YES                  | NO     | YES           | NO      | NO    | Nitrite/sulfite reductase only; CIO bd oxidase  |  | YES              |                |
| <i>Bartonella tamiae</i> (2 strains)                        | Rhizobiales      | YES                  | NO     | YES           | NO      | YES   | <b>NO N metabolism</b> - fragments of proteins related to Nitrite reductase; CIO bd oxidase                         |  | NO               |                |
| <b>minus 4</b>  |                  |                      |        |               |         |       |   |  |                  |                |
| <i>Orientia tsutsugamushi</i> str. Borvong                  | Rickettsiales    | YES                  | YES    | NO            | NO      | NO    | NO N metabolism   |  | YES              |                |
| <i>Rickettsia typhi</i> str. Wilmington                     | Rickettsiales    | YES                  | NO     | YES           | NO      | NO    | NO N metabolism; CIO bd oxidase   |  | YES              |                |
| <i>Bartonella grahamii</i> as4aup                           | Rhizobiales      | YES                  | NO     | NO            | NO      | YES   | NO N metabolism   |  | YES              |                |
| <b>minus 5</b>  |                  |                      |        |               |         |       |   |  |                  |                |
| <i>Bartonella quintana</i> (2 strains)                      | Rhizobiales      | NO                   | NO     | NO            | NO      | YES   | NO N metabolism   |  | YES              |                |
| <i>Bartonella clarridgeiae</i> 73                           | Rhizobiales      | NO                   | NO     | NO            | NO      | YES   | NO N metabolism   |  | YES              |                |
| <i>Bartonella bacilliformis</i> KCS83                       | Rhizobiales      | NO                   | NO     | NO            | NO      | YES   | NO N metabolism   |  | YES              |                |
| <i>Bartonella henselae</i>                                  | Rhizobiales      | NO                   | NO     | NO            | NO      | YES   | NO N metabolism   |  | YES              |                |
| <i>Candidatus Liberibacter asiaticus</i> str. psy62         | Rhizobiales      | NO                   | NO     | NO            | NO      | YES   | NO N metabolism   |  | YES              |                |

**Table S2. Diverse gene clusters for *aa3*-type oxidase in  $\alpha$ -proteobacteria.** The table lists the diverse types of *COX* operons (Fig. 3A). *COX1* proteins recognised as ba3-like\_Oxidase\_I [cd01660] [41] are under the column ba3<sup>^</sup> and correspond to class B [26]. Concatenated operons are framed in blue and connected by a thick line. Incomplete (or ‘dead’ [82]) operons, indicated by the asterisk\*, lack one or more of core subunits *ctaC-E* (Fig. 3A). Functional capacity of the oxidase has been deduced also from biochemical studies [88,89].

| Group and organism   | COX operon <i>aa3</i> -type cytochrome c oxidase |     |            |       |                  | type b                                    | a-b transition | NOTES  |
|--|--|-----|------------|-------|------------------|---|----------------|--|
|  | type a   | a-I | a-II       | a-III | ba3 <sup>^</sup> |   |                |  |
| <b>Beta proteobacteria</b>                                   |  |     |            |       |                  |   |                |  |
| <i>Nitrosomonas europaea</i>                                 |  |     |            |       | yes*             |   | yes            |  |
| <i>Burkholderia xenovorans</i> LB400                         | yes  |     | yes        |       | YES              |   | yes            |  |
| <i>Variovorax paradoxus</i> S110                             | yes  |     | yes        |       |                  |   | yes            |  |
| <i>Methylophilales bacterium</i> HTCC2181                    |  |     |            |       |                  |   | yes            |  |
| <b>Alpha proteobacteria</b>                                  |  |     |            |       |                  |   |                |  |
| <b>Rhodospirillales</b>                                      |  |     |            |       |                  |   |                |  |
| <i>Gluconacetobacter europaeus</i> LMG 18494                 |  | yes |            |       |                  | yes                                       |                |  |
| <i>Gluconacetobacter diazotrophicus</i> PAI 5                | yes  | yes |            |       |                  | yes                                       |                |  |
| <i>Acetobacteraceae bacterium</i> AT-5844                    | yes  |     | yes        | yes   |                  | yes                                       |                |  |
| <i>Roseomonas</i> sp. B-5                                    | yes  | yes |            |       |                  | yes                                       |                | incomplete genome                            |
| <i>Roseomonas cervicalis</i>                                 | yes  |     |            |       |                  | yes                                       |                | incomplete genome                            |
| <i>Tistrella mobilis</i>                                     |  | yes | yes        |       |                  | yes                                       |                |  |
| <i>Caenispirillum salinarum</i>                              |  | yes |            |       | YES              | yes                                       |                |  |
| <i>Magnetospirillum gryphiswaldense</i> MSR-1                | yes - dead*                                      |     |            |       | yes*             | yes                                       |                |  |
| <i>Magnetospirillum magnetotacticum</i>                      |  |     |            |       | yes*             | yes - 1 dead* & 1 complete                |                |  |
| <i>Magnetospirillum magneticum</i> AMB-1                     |  |     |            |       | yes*             | yes - 1 dead* & 1 complete                |                |  |
| <i>Magnetospirillum</i> sp. SO-1                             |  |     |            |       | yes*             | yes - 1 dead* & 1 complete                |                |  |
| <i>Azospirillum brasilense</i>                               | yes  |     | yes        |       |                  | yes                                       |                |  |
| <i>Azospirillum lipoferum</i>                                | yes  |     |            |       |                  | yes - 2                                   |                |  |
| <i>Rhodospirillum centenum</i> SW                            | yes  |     |            |       |                  | yes                                       |                |  |
| <i>Azospirillum amazonense</i>                               |  |     |            |       |                  | yes                                       |                |  |
| <b>Rickettsiales</b>   |  |     |            |       |                  |   |                |  |
| <i>Anaplasma phagocytophilum</i> HZ                          |  |     |            |       |                  | yes - 2 (1 split, 1 dead*)                |                |  |
| <i>Ehrlichia ruminantium</i> str. Welgevonden                |  |     |            |       |                  | yes - 2 (1 split, 1 dead*)                |                |  |
| <b>Rhizobiales</b>   |  |     |            |       |                  |   |                |  |
| <i>Nitrobacter winogradsky</i>                               | yes  |     |            |       |                  | yes - 2                                   |                |  |
| <i>Nitrobacter hamburgensis</i>                              |  | yes |            |       |                  | yes                                       |                |  |
| <i>Bradyrhizobium</i> sp. S23321                             |  |     |            |       | YES              | yes                                       |                | plus other 3 <i>Bradyrhizobium</i> species   |
| <i>Bradyrhizobium</i> sp. WSM471                             | yes - dead*                                      |     | yes        | yes   | YES              | yes                                       |                |  |
| <i>Bradyrhizobium</i> sp. ORS 375                            |  | yes | yes - 2    |       |                  | yes                                       |                |  |
| <i>Bradyrhizobium japonicum</i> USDA 6                       |  |     | yes        |       |                  | yes                                       |                |  |
| <i>Bradyrhizobium</i> sp. ORS 285                            |  |     | yes        |       |                  | yes                                       |                | plus other 3 <i>Bradyrhizobium</i> species   |
| <i>Bradyrhizobium</i> sp. ORS 278                            |  |     |            |       | YES              | yes                                       |                |  |
| <i>Rhodopseudomonas palustris</i> BisA53 (& DX-1)            | yes  |     |            |       |                  | yes                                       |                | plus other 3 <i>Rhodopseudomonas</i> strain  |
| <i>Afipia felis</i>  | yes - dead*                                      |     |            |       |                  | yes                                       |                | incomplete genome                            |
| <i>Afipia birgiae</i>  | yes - dead*                                      |     |            | yes   |                  | yes                                       |                | incomplete genome                            |
| <i>Methylocystis</i> sp. SC2                                 | yes - 2  | yes | yes        | yes   |                  | yes                                       |                |  |
| <i>Methylocystis parvus</i>                                  | yes  | yes |            | yes   |                  | yes                                       |                | incomplete genome                            |
| <i>Methylosinus trichosporium</i>                            | yes  |     |            |       |                  | yes                                       |                |  |
| <i>Microvirga</i>  | yes  |     | yes        | yes   |                  | yes                                       |                |  |
| <i>Methylobacterium extorquens</i> PA1                       |  | yes | yes        |       |                  | yes                                       |                |  |
| <i>Methylobacterium extorquens</i> DM4                       |  | yes |            |       | YES              | yes                                       |                | plus other 3 <i>Methylobacterium</i> species |
| <i>Methylobacterium</i> sp. 4-46                             | yes  |     | yes        |       |                  | yes                                       |                |  |
| <i>Methylobacterium nodulans</i>                             |  | yes | yes        | yes   | YES              | yes                                       |                |  |
| <i>Methylobacterium radiotolerans</i>                        | yes - 2  |     |            |       |                  | yes                                       |                |  |
| <i>Methylobacterium populi</i>                               |  | yes |            |       |                  | yes                                       |                | plus another <i>Methylobacterium</i> species |
| <i>Chelatococcus</i> sp. GW1                                 | yes - dead*                                      |     | yes        | yes   |                  | yes                                       |                | incomplete genome                            |
| <i>Methylocella silvestris</i>                               |  |     | yes        |       |                  | yes                                       |                |  |
| <i>Methyloferula stellata</i>                                |  |     | yes        |       | yes*             | yes                                       |                | incomplete genome                            |
| <i>Beijerinckia indica</i>                                   |  |     |            |       |                  | yes                                       |                |  |
| <i>Starkeya novella</i>                                      | yes  |     |            | yes   |                  | yes                                       |                |  |
| <i>Pseudaminobacter salicylatoxidans</i>                     | yes - dead*                                      |     |            |       |                  | yes                                       |                |  |
| <i>Aurantimonas manganoxydans</i>                            | yes  |     | yes        |       |                  | yes                                       |                |  |
| <i>Pelagibacterium halotolerans</i> B2                       | yes  |     |            |       |                  | yes                                       |                |  |
| <i>Parvibaculum lavamentivorans</i> DS-1                     | yes  |     |            |       |                  | yes                                       |                |  |
| <i>Sinorhizobium meliloti</i> 1021                           | yes  |     |            | yes   |                  | yes                                       |                |  |
| <i>Chelativorans</i> sp. BNC1                                | yes  |     | yes        | yes   |                  | yes                                       |                |  |
| <i>Sinorhizobium medicae</i> WSM419                          | yes  |     | yes        |       |                  | yes                                       |                |  |
| <i>Mesorhizobium opportunistum</i> WSM2075                   |  | yes | yes - 2    |       |                  | yes                                       |                | plus other 3 <i>Mesorhizobium</i> species    |
| <i>Rhizobium leguminosarum</i> bv. trifolii WSM1325          |  | yes | yes        | yes   |                  | yes                                       |                | plus other 5 <i>Rhizobium</i> species        |
| <i>Rhizobium etli</i> CFN 42/mimosae                         |  | yes |            |       | YES              | yes                                       |                | plus other 4 <i>Rhizobium</i> species        |
| <i>Mesorhizobium ciceri</i> biovar <i>biserrulae</i> WSM1271 |  |     | yes        |       |                  | yes                                       |                |  |
| <b>Rhodobacterales</b>                                       |  |     |            |       |                  |   |                |  |
| <i>Rhodobacter sphaeroides</i> 2.4.1                         | yes  |     |            |       |                  | yes                                       |                |  |
| <i>Oceaniovalibus guishaninsula</i>                          |  | yes |            |       |                  | yes                                       |                |  |
| <i>Rhodovulum</i> sp. PH10                                   |  | yes |            |       |                  | yes                                       |                |  |
| <i>Roseobacter litoralis</i>                                 |  |     |            |       |                  | yes                                       |                |  |
| <b>Other groups</b>  |  |     |            |       |                  |   |                |  |
| <i>Brevundimonas subvibrioides</i>                           | yes  |     |            |       |                  | yes                                       |                |  |
| <i>Phenylbacterium zucineum</i> HLK1                         | yes  |     |            |       |                  | yes                                       |                |  |
| <i>Sphingobium</i> sp. SYK-6                                 | yes  |     |            |       |                  | yes                                       |                |  |
| <i>Sphingobium chlorophenolicum</i> L-1                      |  |     | yes, dead* | yes   |                  | yes                                       |                | plus other 4 <i>Sphingobium</i> species      |
| <i>Sphingobium japonicum</i> UT265                           |  |     |            | yes   |                  | yes                                       |                |  |
| <i>Novosphingobium pentaromaticivorans</i>                   |  |     |            | yes   |                  | yes                                       |                |  |
| <i>Novosphingobium</i> sp. PP1Y                              |  |     |            | yes   |                  | yes                                       |                |  |
| <i>Candidatus Pelagibacter</i> sp. IMCC9063                  |  |     |            | yes   |                  | yes                                       |                | plus several marine metagenome organisms     |
| <i>Micavibrio</i>  |  |     |            |       |                  | yes                                       |                |  |
|  | concatenated                                     |     |            |       | ^Type B, [26]    | *incomplete operon, likely non-functional |                |  |

**Table S3 – Phylogenetic distribution of the main characters of COX gene operons.**

We constructed a matrix of 11 independent characters (indicated concisely on top of the columns) that could differentiate the gene sequence of COX subunits in the mitochondria of some protists from the gene sequence of bacterial COX operons (Table S2). The cumulative phenetic analysis indicate that COX operon type a-II of methyloprotoplasts and *Tistrella* (highlighted) share the largest number of characters with COX gene clusters of protist mitochondria (F. Comandatore and C. Bandi, unpublished).

| organism and COX1 length                                    | CHARACTERS-->   | 1        | 2                   | 3           | 4                    | 5             | 6         | 7           | 8               | 9          | 10      | 11                        | total characters vs. mitos |
|---|-----------------|----------|---------------------|-------------|----------------------|---------------|-----------|-------------|-----------------|------------|---------|---------------------------|----------------------------|
|   | COX operon type | COX2COX1 | CtaC-F core COX3 in | E90 in COX3 | Cox3 fused with COX1 | COX3 with 7TM | COX11COX3 | COX4 NOT in | COX4 fused COX1 | COX2 fused | no ctaB | in operon ABC transporter |                            |
| <b>gamma proteobacteria</b>                                 |                 |          |                     |             |                      |               |           |             |                 |            |         |                           |                            |
| <i>Methylococcus capsulatus</i> , 541 aa                    | transition a-b  | 1        | 1                   | 0           | 0                    | 0             | 1         | 0           | 0               | 1          | 1       | 0                         | 5                          |
| <i>Methylobacter marinus</i> , 642 aa                       | a               | 1        | 1                   | 0           | 1                    | 0             | 0         | 0           | 1               | 1          | 1       | 0                         | 6                          |
| <b>beta proteobacteria</b>                                  |                 |          |                     |             |                      |               |           |             |                 |            |         |                           |                            |
| <i>Methylophilales bacterium</i> , 531 aa                   | transition a-b  | 1        | 0                   | 1           | 0                    | 0             | 0         | 0           | 0               | 1          | 0       | 0                         | 3                          |
| <i>Variovorax paradoxus</i> , 662 aa                        | a-II            | 1        | 1                   | 1           | 0                    | 0             | 0         | 0           | 1               | 1          | 1       | 1                         | 6                          |
| <b>alpha proteobacteria</b>                                 |                 |          |                     |             |                      |               |           |             |                 |            |         |                           |                            |
| <i>Acidiphilium cryptum</i>                                 | transition a-b  | 1        | 0                   | 1           | 0                    | 1             | 0         | 0           | 0               | 0          | 0       | 0                         | 3                          |
| <i>Rhodospseudomonas palustris HaA2</i>                     | a               | 1        | 1                   | 1           | 0                    | 1             | 0         | 0           | 0               | 0          | 0       | 0                         | 4                          |
| <i>Methylocystis</i> sp. SC2, 636 aa                        | a-II            | 1        | 1                   | 1           | 0                    | 0             | 1         | 0           | 1               | 1          | 1       | 0                         | 7                          |
| <i>Mesorhizobium opportunisticum</i> , 658 aa               | a-II            | 1        | 1                   | 1           | 0                    | 0             | 0         | 0           | 1               | 1          | 1       | 1                         | 7                          |
| <i>Microvirga</i> sp. WSM355B, 670 aa                       | a-II            | 1        | 1                   | 1           | 0                    | 0             | 0         | 0           | 1               | 1          | 1       | 0                         | 7                          |
| <i>Methylocella sylvestris</i> , 650 aa                     | a-II            | 1        | 1                   | 1           | 0                    | 0             | 0         | 0           | 1               | 1          | 1       | 1                         | 7                          |
| <i>Methylobacterium extorquens PA1</i> , 638 aa             | a-II            | 1        | 1                   | 1           | 0                    | 0             | 1         | 1           | 1               | 1          | 1       | 1                         | 9                          |
| <i>Rhizobium leguminosarum</i> bv. <i>Trifolii</i> , 637 aa | a-II            | 1        | 1                   | 1           | 0                    | 0             | 1         | 1           | 1               | 1          | 1       | 0                         | 8                          |
| <i>Tistrella mobilis</i> , 839 aa                           | a-II            | 1        | 1                   | 1           | 1                    | 0             | 0         | 0           | 1               | 1          | 1       | 0                         | 7                          |
| <i>Azospirillum brasilense</i> , 645 aa                     | a-II            | 1        | 1                   | 1           | 0                    | 0             | 0         | 0           | 1               | 1          | 1       | 0                         | 6                          |
| <i>Beijerinckia indica</i> subsp. <i>indica</i> , 555 aa    | b               | 1        | 0                   | 1           | 0                    | 1             | 1         | 1           | 0               | 0          | 0       | 0                         | 5                          |
| <i>Micavibrio aeruginovorans</i> , 559 AA                   | b               | 1        | 0                   | 1           | 0                    | 1             | 1         | 1           | 0               | 0          | 0       | 0                         | 5                          |
| <i>Magnetospirillum magneticum</i> AMB-1                    | b               | 1        | 0                   | 1           | 0                    | 1             | 1         | 1           | 0               | 0          | 0       | 0                         | 5                          |
| <i>Margaritula maris</i>                                    | b               | 1        | 0                   | 1           | 0                    | 1             | 1         | 1           | 0               | 0          | 0       | 0                         | 5                          |
| <i>Rhoseobacter litoralis</i>                               | b               | 0        | 0                   | 1           | 0                    | 1             | 1         | 0           | 0               | 0          | 0       | 0                         | 3                          |
| <i>Wolbachia</i> wRI  | b               | 1        | 0                   | 1           | 0                    | 1             | 0         | 1           | 0               | 0          | 0       | 0                         | 4                          |
| <i>Rickettsia felis</i>                                     | b               | 0        | 0                   | 1           | 0                    | 1             | 0         | 1           | 0               | 0          | 0       | 0                         | 3                          |
| <b>eukaryotes (mitochondria)</b>                            |                 |          |                     |             |                      |               |           |             |                 |            |         |                           |                            |
| <i>Dictyostelium</i>  |                 | 1        | 0                   | 1           | 0                    | 1             | 0         | 1           | 0               | 1          | 1       | 0                         | 6                          |
| <i>Jakobida</i>   |                 | 1        | 0                   | 1           | 0                    | 1             | 1         | 1           | 0               | 0          | 1       | 1                         | 7                          |
| <i>Tetrahymena</i>  |                 | 1        | 0                   | 1           | 0                    | 0             | 0         | 0           | 1               | 0          | 1       | 0                         | 4                          |
| <i>Oxytricha</i>  |                 | 1        | 1                   | 1           | 1                    | 1             | 0         | 0           | 1               | 1          | 1       | 0                         | 8                          |
|   | CHARACTERS      | 1        | 4                   | 6           | 3                    | 5             | 7         | 8           | 8               | 9          |         | 11                        |                            |

**Table S4 – Conserved phospholipid binding sites in COX3 proteins.**

The alignment in Fig. 4A was enlarged and the residues corresponding to the PL-binding sites and E90 (close to O2 entry in beef COX3 [60]) were considered conserved when producing positive substitutions [83] (bold amino acid symbols in white background). Other substitutions are highlighted in pale brown while identities are identified as yes. Organisms are abbreviated as in Fig. 4.

| in phylogenetic/phenetic order |                |              |             | PL binding sites (beef numbering) |     |     |     |     |           |      |      |        |         |      |      |      |      |      |      |           | PE1       | PL binding total* |                   |
|--------------------------------|----------------|--------------|-------------|-----------------------------------|-----|-----|-----|-----|-----------|------|------|--------|---------|------|------|------|------|------|------|-----------|-----------|-------------------|-------------------|
| Organism and protein           | protein        | Accession    | length (aa) | O2                                | PG1 |     |     |     | PG1 & PG2 |      |      |        | PG2     |      |      |      | PE2  |      |      | G82 & M83 | total*    |                   |                   |
|                                |                |              |             | E90                               | W57 | W58 | E64 | F86 | H71       | T66  | F214 | R221   | F/V233  | G234 | Y241 | H231 | Y181 | Y182 | I185 |           |           | F198              | F203              |
| <i>Methylocystis</i> sp. SC2   | COX1-3         | YP_006591291 | 827         | L                                 | yes | yes | N   | G   | L         | P    | C    | A      | T       | S    | L    | P    | yes  | A    | gap  | A         | A         | 0                 | <b>3</b>          |
| <i>Rhodo_palu_BisA53</i>       | COX1-3         | YP_782773    | 841         | yes                               | P   | G   | A   | G   | R         | Y    | gap  | S      | A       | M    | I    | T    | M    | L    | G    | yes       | 0         | <b>2 &amp; 2</b>  |                   |
| <i>Methylocystis</i> sp. SC2   | COX3a          | YP_006592539 | 236         | V                                 | M   | E   | yes | V   | yes       | gap  | A    | K      | T       | R    | yes  | L    | W    | R    | gap  | Y         | V         | 0                 | <b>3 &amp; 3</b>  |
| <i>Variovorax</i> par EPS      | COX3b          | WP_018905232 | 234         | D                                 | gap | gap | gap | yes | Y         | gap  | yes  | K      | S       | I    | yes  | N    | F    | S    | gap  | yes       | yes       | 0                 | <b>5 &amp; 4</b>  |
| <i>Methylocystis</i> sp. SC2   | COX3           | YP_006592793 | 213         | yes                               | gap | gap | gap | I   | yes       | G    | L    | K      | N       | S    | yes  | R    | W    | R    | A    | yes       | L         | 0                 | <b>3 &amp; 3</b>  |
| <i>Roseomonas</i> sp. B5       | COX3           | WP_019459297 | 217         | yes                               | gap | gap | A   | yes | yes       | E    | V    | Q      | M       | P    | yes  | Q    | yes  | H    | gap  | Y         | L         | 1                 | <b>5 &amp; 4</b>  |
| <i>Gluconacetobacter</i> europ | COX3           | WP_019086564 | 210         | yes                               | gap | gap | yes | yes | yes       | P    | L    | yes    | P       | A    | yes  | R    | W    | H    | A    | Y         | L         | 2                 | <b>7 &amp; 4</b>  |
| <i>Variovorax</i> par S110     | COX3           | YP_002942615 | 201         | yes                               | gap | gap | gap | L   | yes       | gap  | D    | L      | V       | D    | yes  | R    | F    | A    | W    | W         | L         | 1                 | <b>3 &amp; 4</b>  |
| <i>Methylocystis</i> sp. SC2   | COX3           | YP_006593114 | 242         | yes                               | gap | gap | gap | yes | Y         | gap  | D    | L      | S       | D    | yes  | R    | F    | A    | W    | W         | L         | 1                 | <b>3 &amp; 4</b>  |
| <i>Azospirillum</i> brasilense | COX3           | YP_005030787 | 202         | yes                               | gap | gap | gap | yes | yes       | gap  | D    | L      | Q       | D    | yes  | R    | F    | T    | W    | W         | yes       | 1                 | <b>5 &amp; 3</b>  |
| <i>Nitrosococcus</i> oceani    | COX3           | YP_345015    | 288         | yes                               | yes | F   | yes | yes | gap       | E    | M    | yes    | yes (F) | A    | yes  | yes  | yes  | G    | D    | Y         | yes       | 2                 | <b>11 &amp; 3</b> |
| <i>Methylophi</i> bac          | COX3           | WP_008107990 | 299         | yes                               | yes | F   | yes | yes | K         | R    | M    | yes    | yes (F) | A    | yes  | D    | yes  | M    | N    | Y         | yes       | 2                 | <b>10 &amp; 3</b> |
| <i>Thermus</i> thermophilus    | COX1-3         | YP_143578    | 791         | yes                               | yes | A   | yes | yes | yes       | S    | V    | L      | gap     | T    | yes  | yes  | F    | yes  | H    | W         | yes       | 2                 | <b>10 &amp; 3</b> |
| mitochondria                   |                |              |             |                                   |     |     |     |     |           |      |      |        |         |      |      |      |      |      |      |           |           |                   |                   |
| <i>Plasmodium</i> yoelii       | COX3           | XP_729225    | 263         | yes                               | F   | V   | yes | T   | gap       | L    | yes  | D      | yes (V) | V    | yes  | S    | L    | G    | gap  | H         | T         | 2                 | <b>6 &amp; 2</b>  |
| <i>Monoeuroples</i> minuta     | COX1-3         | ACX30949     | 1203        | S                                 | gap | F   | L   | I   | gap       | yes  | Y    | L      | I       | T    | V    | F    | F    | D    | L    | yes       | 1         | <b>4 &amp; 5</b>  |                   |
| <i>Oxytricha</i> trifallax     | COX1-3         | AEV66698     | 1331        | S                                 | gap | F   | Y   | gap | F         | I    | Y    | yes    | Y       | N    | yes  | W    | F    | D    | yes  | yes       | N         | 1                 | <b>5 &amp; 6</b>  |
| alpha, sister group of mitos   |                |              |             |                                   |     |     |     |     |           |      |      |        |         |      |      |      |      |      |      |           |           |                   |                   |
| <i>Beijerinckia</i> indica     | COX3           | YP_001831864 | 287         | yes                               | yes | yes | yes | yes | yes       | E    | yes  | yes    | yes     | yes  | yes  | Q    | F    | A    | D    | yes       | yes       | 2                 | <b>14 &amp; 3</b> |
| <i>Caenispirillum</i>          | COX3           | WP_009539949 | 267         | yes                               | yes | yes | yes | yes | yes       | M    | yes  | yes    | S       | I    | W    | R    | yes  | G    | A    | Y         | yes       | 2                 | <b>11 &amp; 3</b> |
| <i>Ricket</i> prow_Rp22        | COX3           | YP_005998599 | 278         | yes                               | yes | yes | yes | yes | yes       | I    | yes  | yes    | L       | yes  | yes  | G    | yes  | yes  | E    | Y         | yes       | 2                 | <b>14 &amp; 1</b> |
| <i>Wolbachia</i> Dro_sim       | COX3           | YP_007886075 | 275         | yes                               | yes | yes | yes | yes | yes       | I    | yes  | yes    | L       | C    | yes  | D    | yes  | yes  | F    | Y         | yes       | 1                 | <b>11 &amp; 1</b> |
| <i>Paracoccus</i>              | COX3           | YP_918077    | 273         | yes                               | yes | yes | yes | yes | yes       | E    | yes  | yes    | yes     | yes  | yes  | Q    | yes  | S    | L    | Y         | yes       | 2                 | <b>14 &amp; 4</b> |
| <i>Roseobacter</i>             | COX3           | YP_004690241 | 273         | yes                               | yes | yes | yes | yes | yes       | K    | yes  | yes    | yes     | yes  | yes  | K    | yes  | S    | L    | yes       | yes       | 2                 | <b>15 &amp; 2</b> |
| <i>Tistrella</i>               | COX3           | YP_006372231 | 269         | yes                               | yes | yes | yes | yes | yes       | Q    | yes  | yes    | I       | yes  | yes  | K    | yes  | S    | T    | Y         | yes       | 2                 | <b>13 &amp; 2</b> |
| <i>Acidiphilium</i>            | COX3           | YP_001235282 | 283         | yes                               | yes | yes | yes | yes | yes       | R    | yes  | yes    | yes     | yes  | yes  | R    | yes  | T    | P    | yes       | yes       | 2                 | <b>15 &amp; 1</b> |
| sister group2: mitochondria    |                |              |             |                                   |     |     |     |     |           |      |      |        |         |      |      |      |      |      |      |           |           |                   |                   |
| <i>Dictyostelium</i>           | COX3           | NP_050071    | 435         | yes                               | yes | yes | D   | yes | yes       | L    | T    | V      | Q       | I    | yes  | N    | yes  | yes  | gap  | yes       | yes       | 2                 | <b>11 &amp; 1</b> |
| <i>Acanthamoeba</i>            | COX3           | NP_042532    | 329         | yes                               | yes | yes | yes | yes | yes       | yes  | yes  | Q      | yes (F) | yes  | yes  | yes  | yes  | G    | gap  | H         | L         | 1                 | <b>13 &amp; 1</b> |
| <i>Andalucia</i>               | COX3           | YP_007890482 | 267         | yes                               | yes | yes | yes | yes | yes       | yes  | yes  | yes    | yes     | yes  | yes  | yes  | G    | gap  | H    | yes       | 1         | <b>15 &amp; 1</b> |                   |
| <i>Bos taurus</i> (beef)       | COX3           | AAZ16929     | 260         | yes                               | yes | yes | yes | yes | yes       | yes  | yes  | yes    | yes     | yes  | yes  | yes  | yes  | yes  | yes  | yes       | yes       | 2                 | <b>19</b>         |
| Organism and protein           | protein operon | aa           | E90         | W57                               | W58 | E64 | F86 | H71 | T66       | F214 | R221 | F/V233 | G234    | Y241 | H231 | Y181 | Y182 | I185 | F198 | F203      | G82 & M83 | total*            |                   |