Subsystem: TCA Cycle

Olga Vassieva[1] and Rick Stevens[2]
[1]FIG, [2]Argonne National Laboratory and University of Chicago

- Tricarboxylic acid cycle (TCA) oxidizes acetyl-CoA to CO_2 when functions in its full capacity. It supplies NADH for oxidative phosphorylation in the respiratory chain and thus has a major role in energy metabolism. It is also a source of three essential precursor metabolites. Many organisms contain partial TCA, functioning mainly as a source of precursor metabolites. Specifically, 2-ketoglutarate, oxaloacetate, and succinyl-CoA are starting points for the biosynthesis of glutamate, aspartate, and porphyrin,respectively

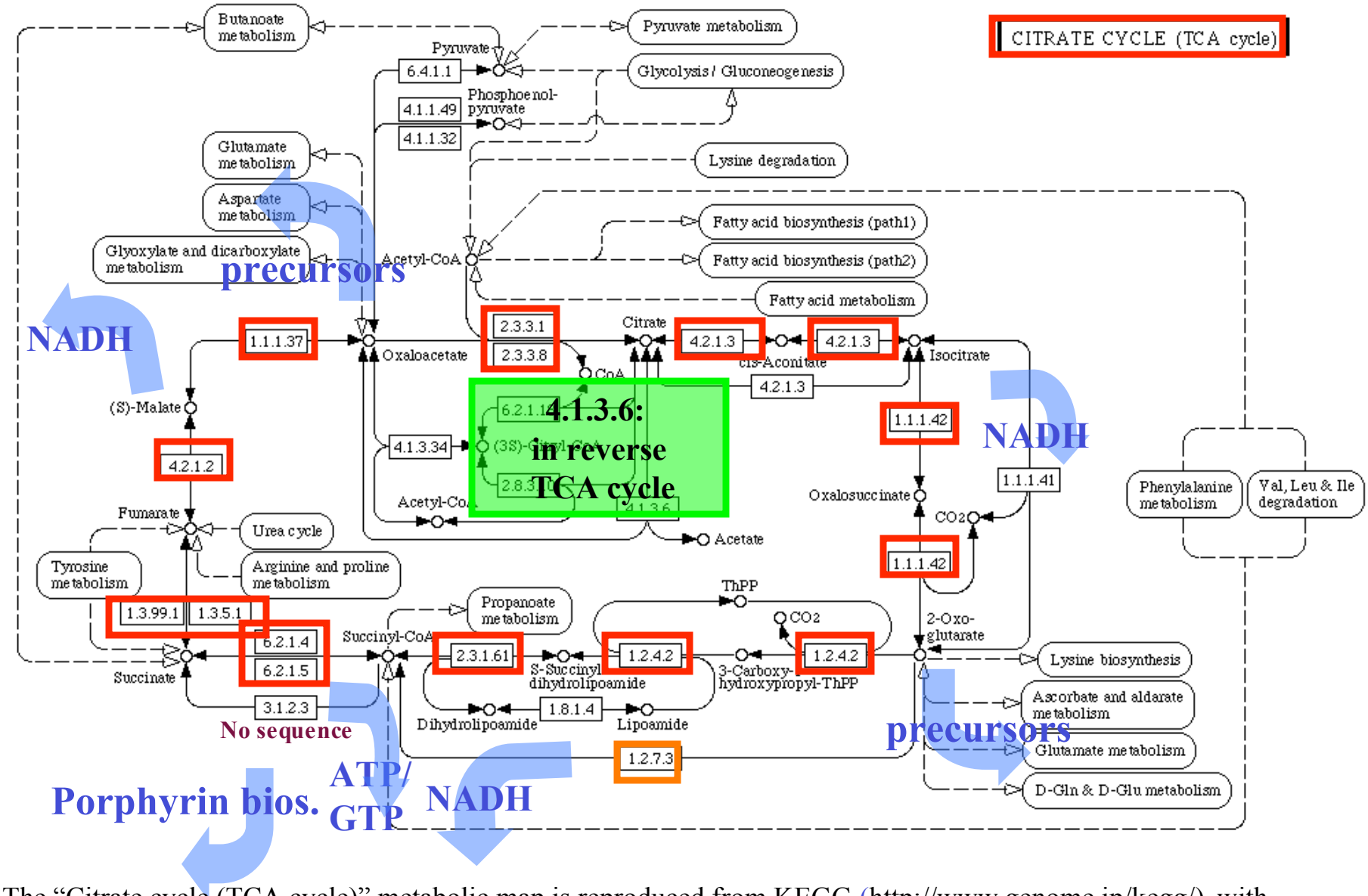
- Green bacteria and archaea assimilate CO_2 by employing reducing power (in the form of reduced ferredoxin) to reverse the normal direction of carbon flow through the TCA cycle and to produce 6 metabolites. "Reductive carboxylic acid cycle" (also called the reductive tricarboxylic acid cycle) appeared to be the sole CO_2 assimilation pathway in *Chlorobium thiosulfatophilum* (Evans et al. 1966).

- This Subsystem has been annotated across 292 genomes. About 120 of them possess complete set of functional roles. Approximately the same number of organisms appear to contain other physiologically relevant variants of this subsystem. Some of them are discussed below.

List of Functional roles

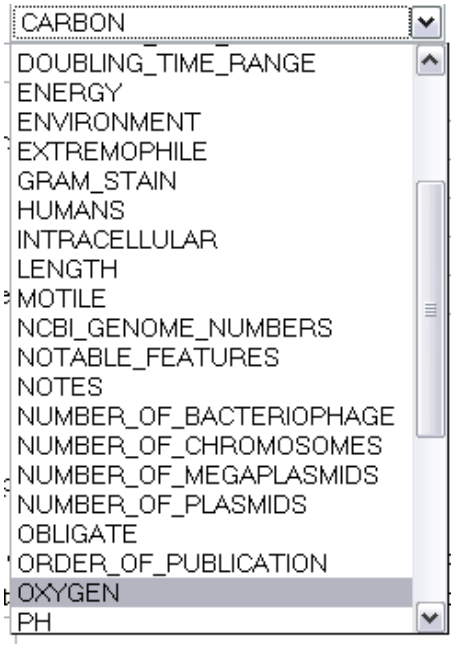
Column	Abbrev	Fun
1	gltA	Citrate synthase (si) (EC 2.3.3.1)
2	acnB	Aconitate hydratase 2 (EC 4.2.1.3)
3	acnA	Aconitate hydratase (EC 4.2.1.3)
4	icd	Isocitrate dehydrogenase [NADP] (EC 1.1.1.42)
5	sucA	2-oxoglutarate dehydrogenase E1 component (EC 1.2.4.2)
6	sucB	2-oxoglutarate dehydrogenase E2 component (EC 2.3.1.61)
7	lpdA	Dihydrolipoamide dehydrogenase (EC 1.8.1.4)
8	sucD	Succinyl-CoA ligase [ADP-forming] alpha chain (EC 6.2.1.5)
9	sucC	Succinyl-CoA ligase [ADP-forming] beta chain (EC 6.2.1.5)
10	sdhB	Succinate dehydrogenase iron-sulfur protein (EC 1.3.99.1)
11	sdhA	Succinate dehydrogenase flavoprotein subunit (EC 1.3.99.1)
12	fumA	Fumarate hydratase class I, aerobic (EC 4.2.1.2)
13	fumC	Fumarate hydratase class II (EC 4.2.1.2)
14	mdh	Malate dehydrogenase (EC 1.1.1.37)
15	mqo	Malate:quinone oxidoreductase (EC 1.1.99.16)
16	IDHI	Isocitrate dehydrogenase [NAD] subunit I, mitochondrial precursor (EC 1.1.1.41)
17	IDHII	Isocitrate dehydrogenase [NAD] subunit II, mitochondrial precursor (EC 1.1.1.41)
18	ICD3g	Isocitrate dehydrogenase 3 subunit gamma, mitochondrial precursor (EC 1.1.1.41)
19	ICD3b	Isocitrate dehydrogenase 3 subunit beta, mitochondrial precursor (EC 1.1.1.41)
20	ICD3a	Isocitrate dehydrogenase 3 subunit alpha, mitochondrial precursor (EC 1.1.1.41)
21	IDHP	putative isocitrate dehydrogenase (EC 1.1.1.41)
22	ICDX	NAD-dependent isocitrate dehydrogenase precursor (EC 1.1.1.41)
23	SUCLG1	Succinyl-CoA ligase [GDP-forming] alpha chain (EC 6.2.1.4)
24	SUCLG2	Succinyl-CoA ligase [GDP-forming] beta chain (EC 6.2.1.4)
25	FumB(N)	Tartrate dehydratase alpha subunit/Fumarate hydratase class I, N-terminal domain
26	AHX	Aconitate hydratase X, predicted
27	OGSg	2-oxoglutarate oxidoreductase, gamma subunit (EC 1.2.7.3)
28	OCSb	2-oxoglutarate oxidoreductase, beta subunit (EC 1.2.7.3)
29	OGSa	2-oxoglutarate oxidoreductase, alpha subunit (EC 1.2.7.3)
30	OGSd	2-oxoglutarate oxidoreductase, delta subunit, putative (EC 1.2.7.3)
31	FumHan	Fumarate hydratase class I, anaerobic (EC 4.2.1.2)
32	icdNAD	Isocitrate dehydrogenase [NAD] (EC 1.1.1.41)
33	ATPCS	ATP citrate synthase (EC 2.3.3.8)
34	ATPCS1	citryl-CoA lyase (EC 4.2.3.34)
35	ATPCS2	citrate-CoA ligase (EC 6.2.1.18)
36	SCoAH	succinyl-CoA hydrolase (EC 3.1.2.3)
37	C3SL	citrate (pro-3S)-lyase (EC 4.1.3.6)
38	CCoAT	citrate CoA-transferase (EC 2.8.3.10)
39	CyCoAL	citryl-CoA lyase (EC 4.1.3.34)

TCA Cycle: Subsystem diagram



The “Citrate cycle (TCA cycle)” metabolic map is reproduced from KEGG (http://www.genome.jp/kegg/) with modifications. Functional roles are indicated by the corresponding EC numbers, blue arrows show output of NADH and essential metabolites produced in TCA.

Subsystem Spreadsheet (a fragment)



Rows are coloured by one of organism's attributes available in SEED: oxygen. Orange - aerobic, blue - anaerobic, dark magenta - facultatively aerobic, light magenta - facultatively anaerobic. Comments in red **(I, II, III)** are discussed in detail in the next slides.

Organism	Variant Code	gltA	*AcoH	*icd	sucA	sucB	lpdA	sucD	sucC	sdhB	sdhA	*Fum	*MD
Escherichia coli K12 [B]	1	712	118-2, 1263-3	1121-4	718	719	116	721	720	4066, 716	715, 4067	1598-12, 1597-13	3180-14, 3509-14, 2184-15
Escherichia coli O157:H7 EDL933 [B]	1	744	122-2, 2229-3	1658-4	750	751	120	753	752	748, 5096	747, 5097	2355-12, 757-12, 5063-12, 2354-13	4112-14, 3098-15
Erwinia carotovora atroseptica [B]	1	575, 3100	1674-2, 1496-3	456-4	570	3509, 569	1683	890	3506, 891	3787, 571	572, 3788	4202-12, 928-13	4473-14, 515-15
Bacillus subtilis subsp. subtilis str. 168 [B]	1	2917, 944	1804-3	2916-4	1942	1941	1463, 809, 2410	1612	1611	2846	2847	3310-13	2915-14
Staphylococcus aureus NCTC 8325 [B]	0	349	850-3	347-4, 348-4	1540	1539	2173, 2264		770	471, 479	470, 703	1316-13	2568-15, 2320-15
Staphylococcus aureus subsp. aureus MW2 [B]	1	1639	1237-3	1638-4	1303	1302	1471, 979	1129	1128	1032	1031	1792-13	2286-15, 2526-15
Chlamydophila pneumoniae J138 [B]	123					375	831	971	970	788	787	1009-13	1024-14
Chlamydophila pneumoniae TW-183 [B]	123					387	858	1006	1005	814	813	1047-13	1063-14
Chlorobium tepidum TLS [B]	2	1807	537-2	349-4					1074, 378	2232, 2010	2011, 2233	821-12	1486-14
Chloroflexus aurantiacus [B]	0	2227		336-4	2125	1607, 2124	1042, 1800	2347	2268	3916	3963	3239-13	1347-14, 1573-14
Chromobacterium violaceum ATCC 12472 [B]	1	1070	2470-2, 1121-3, 2054-3	3664-4	1071	1072	1074, 528	1076	1075	1068, 3368	1067, 3369	3476-12, 1120-13	1062-14
Organism	**Variant Code**	**gltA**	***AcoH**	***icd**	**sucA**	**sucB**	**lpdA**	**sucD**	**sucC**	**sdhB**	**sdhA**	***Fum**	***MD**
Clostridium acetobutylicum ATCC 824 [B]	0		1136-3	1137-4								3220-12, 444-13	734-14
Clostridium botulinum ATCC 3502 [B]	0		370-3	371-4								1470-12, 1469-12	

II. oxoglutarate as a starting metabolite: Utilization of Glutamate,lysine?

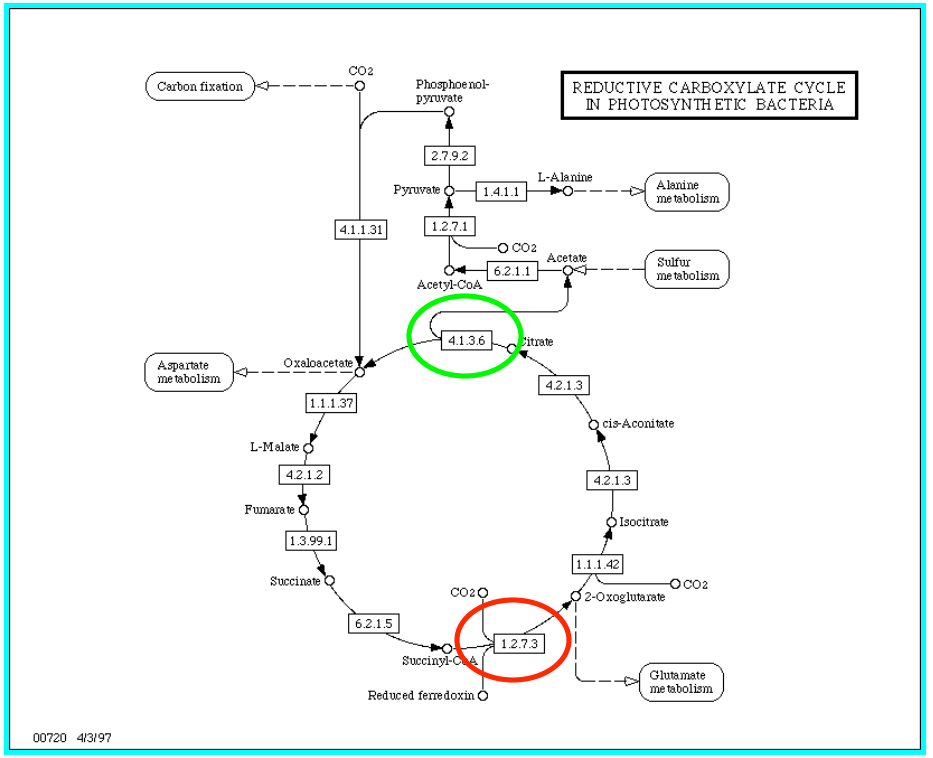
I.Reversed TCA

III. Missing oxoglutarate-Fumarate branch and citrate synthase (Citrate lyase is present): Utilization of citrate as a starting metabolite

Open questions, comments, conjectures

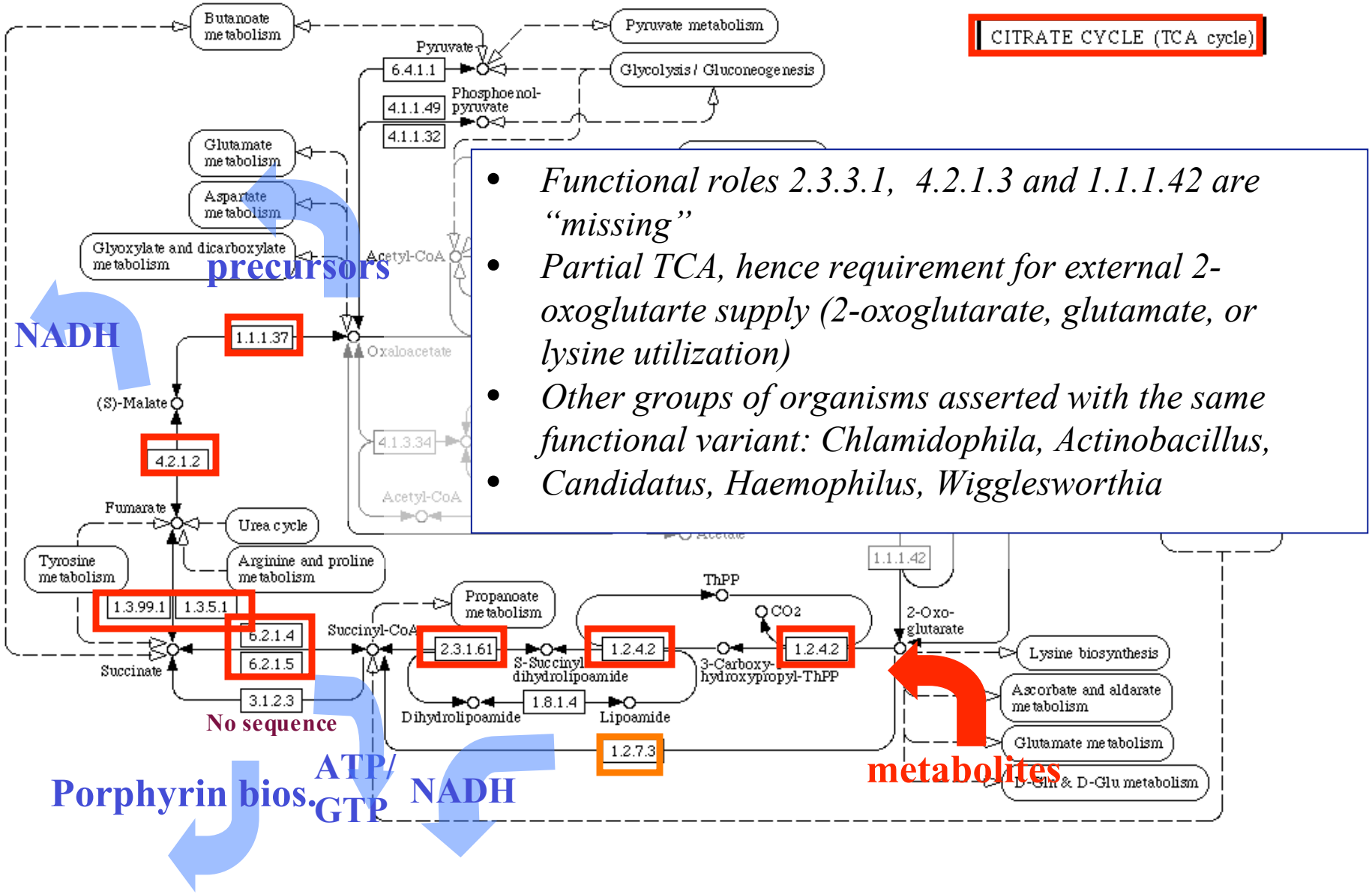
I. Reverse TCA Cycle in Chlorobium

- *2-oxoglutarate synthase (EC 1.2.7.3) is essential, but genes for the 2 out of 4 subunits of this complex cannot be identified in two Chlorobia genomes.*
- *The complete set of 2-oxoglutarate synthase subunits in Archeoglobus is shown for comparison*
- *citrate aldolase (EC 4.1.3.6) is essential in reverse TCA cycle in Chlorobium tepidum*
- *See also the Notes section on the TCA Cycle subsystem page in SEED*



Organism	Variant Code	gltA	*AcoH	*icd	sucA	sucB	lpdA	sucD	sucC	sdhB	sdhA	*Fum	*MD	OGSg	OCSb	OGSa	OGSd	FumHan	ATPCL
Chlorobium tepidum TLS [B]	2	1807	537-2	349-4				267	1074, 378	2010, 2232	2011, 2233	821-12	1486-14	?	162	163	?		1073
Archaeoglobus fulgidus DSM 4304 [A]	2	1330	55-26	640-4				1529, 2171	1530, 2172	675	1453, 674	1086-12, 1087-12	848-14	466	463	464	465		

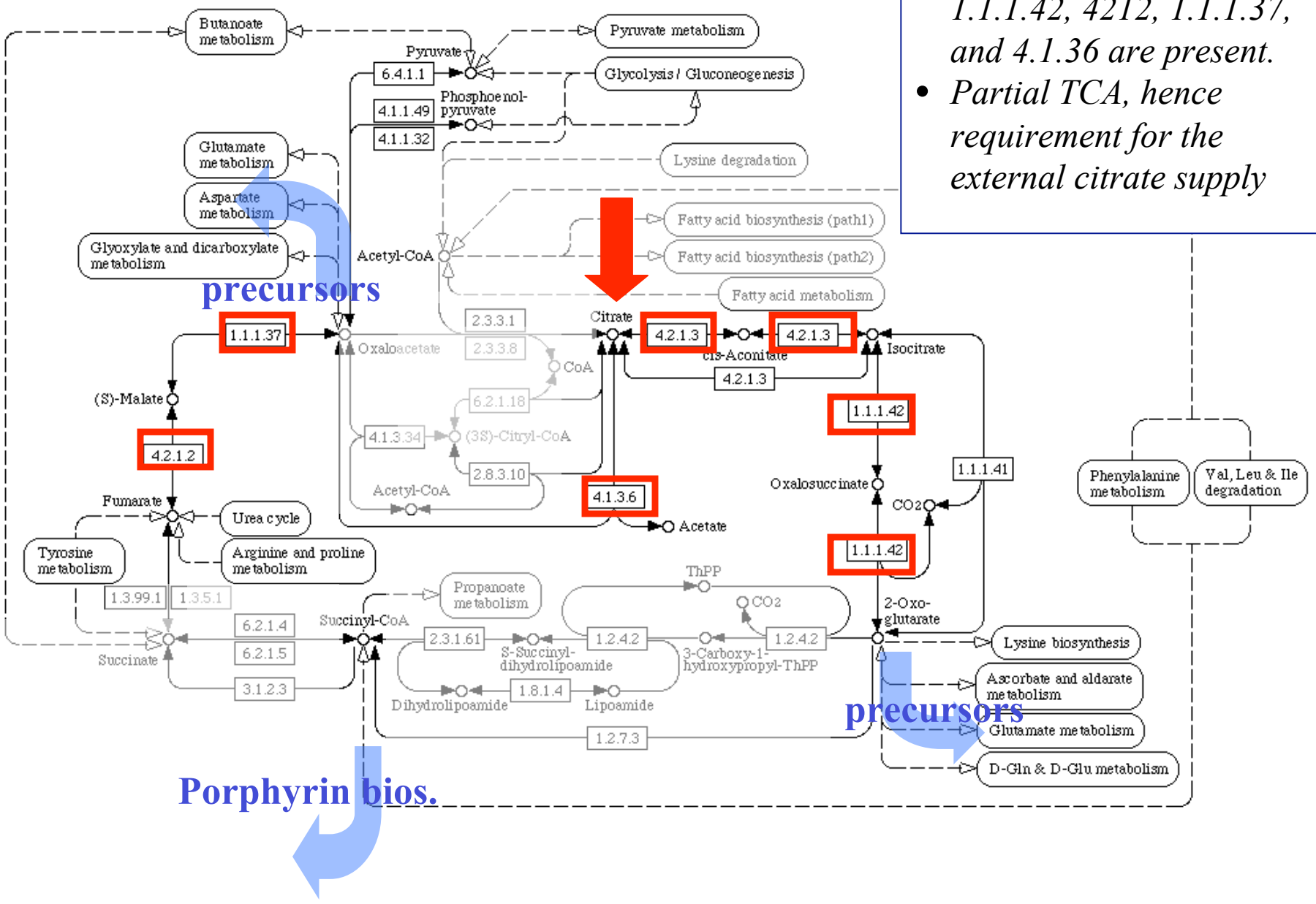
II. "Chlamidia-type" TCA - functional variant 123



- *Functional roles 2.3.3.1, 4.2.1.3 and 1.1.1.42 are "missing"*
- *Partial TCA, hence requirement for external 2-oxoglutarte supply (2-oxoglutarate, glutamate, or lysine utilization)*
- *Other groups of organisms asserted with the same functional variant: Chlamidophila, Actinobacillus,*
- *Candidatus, Haemophilus, Wigglesworthia*

II. "Clostridium-type" TCA - functional variant 145

- *Only roles 4.2.1.3, 1.1.1.42, 4212, 1.1.1.37, and 4.1.36 are present.*
- *Partial TCA, hence requirement for the external citrate supply*



III. Functional coupling of the TCA genes

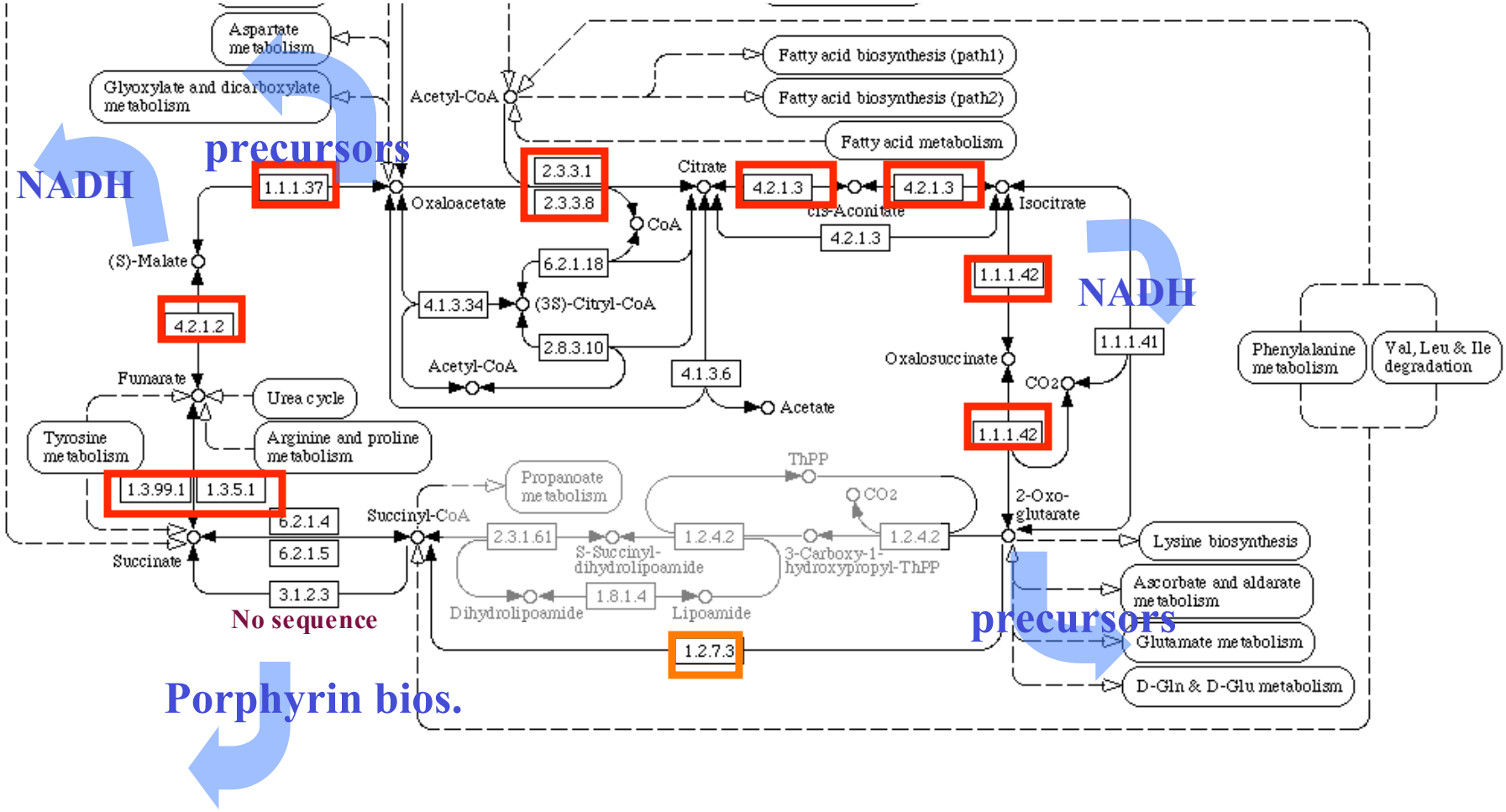
- *A fragment of subsystem spreadsheet is shown*
- *Matching colors highlight genes that occur close to each other on the chromosome*
- *In different genomes orthologous genes are clustered in various combinations*
- *Genes encoding subunits of enzymatic complexes form very conservative clusters*
- *Aconitase homologs often cluster with LpdA*

Organism	Variant Code	gltA	*AcoH	*icd	sucA	sucB	lpdA	sucD	sucC	sdhB	sdhA	*Fum	*MD
Desulfuromonas acetoxidans [B]	1	408, 458	569-2, 1761-2, 2466-3, 107-3, 2660-3	2869-4	2541	2542	1183, 2014, 2543, 2676, 400	3228, 3834	2052, 1293	2210, 911, 3	2209, 910	44-12, 790-12, 2347-13, 1243-13	3088-14
Synechococcus elongatus PCC 7942 [B]		1958	250-2	1035-4	IV. Missing oxoglutarate-succinate branch					857	1931	350-13	
Synechococcus sp. WH 8102 [B]	11	2266	2491-2	163-4						585	586	632-13	1679-15
Synechocystis sp. PCC 6803 [B]	2	2293	2978-2	256-4	Missing oxoglutarate dehydr		1801	365		1192, 2564	940	2231-13	2501-14
Escherichia coli O157:H7 [B]	1	816	209-2, 1883-3	1653-4	822	823	207	825	824	820, 5113	5114, 819	2347-12, 5082-12, 828-12, 2346-13	4083-14, 3093-15
Exiguobacterium sp. 255-15 [B]	1	1287	1517-3	1288-4	2031	1593, 2033	1596, 1779, 2676	2462	2463	1340	1341	1108-13	1289-14, 145-15
Ferroplasma acidarmanus [A]	2	1245, 1919	1018-3	1622-4	Missing oxoglutarate dehydrogenase		1385, 1805	1027	1026	1033	1034		1035-14

Comment IV is discussed in the next slide

IV.”Synechococcus and Synechocystis”- type TCA - functional variants 2 and 11

- *The absence of 2-oxoglutarate dehydrogenase complex is absent (variant code 2) is characteristic for anaerobic growth on glucose and for photoautotrophs. Oxaloacetate is the primary starting metabolite, from which oppositely directed TCA branches operate. Other organisms asserted with this functional variant: Nostoc, Sulfolobus, Aquifex, Bacteroides,Clorobium, Ferroplasma,*
- *The absence of 1.2.4.2 can be compensated by 1.2.7.3. (Wolinella, Campylobacter, Archaeoglobus, Chlorobium(?), Methanococcus, Methanothermobacter, Desulfovibrio, etc)*
- *6.2.1.4 is missing in Synechococcus, Prochlorococcus, Gloeobacter, Helicobacter: these organisms have other active routes for production of Succinyl –CoA required for porphyrin biosynthesis.*



Comments, conclusions

- Only a few functional variants of the TCA cycle (out out many asserted in SEED) were discussed here
- Potential sources of mistakes in the interpretation of functional variants of a subsystem:
  - "Overannotation" of paralogous gene families
  - Miscalled and uncalled ORFs: short (less than 100 amino acids) ORFs are often missed by automatic ORF calling software. For example, delta subunit of 2-oxoglutarate oxidoreductase, the shortest protein of this complex, has not been called in the majority of genomes where three other subunits are clearly present
  - In organisms with unknown physiology it is a nontrivial task to distinguish between the cases of "missing genes" due to (yet unidentified) non-orthologous gene displacements, and those where functional role is lacking for physiological reasons
- Lipoamide dehydrogenase encoded by *lpdA* is the shared component of the 2-oxoglutarate dehydrogenase and pyruvate dehydrogenase complexes. Its sole presence in a genome (and in the subsystem spreadsheet) can be confusing. Analysis of the corresponding gene cluster often helps identify the actual complex LpdA belongs to in a particular organism.
- Some function roles have no sequences associated with them in any organism (globally "missing"). In genomes with "missing" functional role 6.2.1.4, the presence of the potentially complementing 3.2.1.3 can be only inferred, but not established, since no corresponding sequence has been identified