

Title	代謝系遺伝子の分子系統樹比較によるゲノム構造解析
Author(s)	野口, 大毅
Citation	
Issue Date	2000-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/653
Rights	
Description	Supervisor:小長谷 明彦, 知識科学研究科, 修士

修 士 論 文

代謝系遺伝子の分子系統樹比較によるゲノム構造解析

指導教官 小長谷 明彦 教授

北陸先端科学技術大学院大学
知識科学研究科知識システム基礎学専攻

野口 大毅

目次

1 序章	1
1.1 背景	1
1.2 目的	1
2 分子系統学における前提と諸問題	4
2.1 分子進化	4
2.1.1 進化概説	4
2.1.2 分子時計	5
2.1.3 自然淘汰説と分子進化の中立説	6
2.1.4 中立突然変異と進化速度	7
2.1.5 遺伝子の多様性	8
2.1.6 中立進化の検証	9
2.1.7 機能遺伝子の分子進化と系統	10
2.2 配列比較の方法	12
2.2.1 アラインメントの概念	12
2.2.2 アラインメントのアルゴリズム	13
2.2.3 アミノ酸の突然変異	15
3 代謝経路の構造及び進化について	17
3.1 代謝	17
3.2 栄養物質の分解	18
3.3 酸化的リン酸化	20
3.3.1 クエン酸回路	20
3.3.2 エネルギー運搬物質	21
3.4 光合成回路	22

4 既存のデータベースから得られた知見	29
4.1 データの入手	29
4.2 クエン酸回路の形成に関する仮説	33
4.3 クエン酸回路遺伝子の分類	35
5 分子進化系統樹による分類結果の検証	36
5.1 アミノ酸由来	37
5.2 光合成由来に関しての系統樹	40
5.3 炭水化物分解系由来	46
5.4 クエン酸回路固有な酵素群	49
5.5 まとめ	50
6 考察	52
6.1 光合成由来遺伝子の系統樹比較	54
6.2 アミノ酸由来遺伝子の系統樹比較	56
6.3 アライメント比較	59
7 結語	61
謝辞	61
A ENZYME データベースから得られた、PATHWAY マップの詳細	65
B 本論文で対象とした酵素データ	67

目 次

2.1	ヘモグロビンの分子時計	5
2.2	中立進化と、適応進化	9
2.3	機能遺伝子の進化	10
2.4	アライメント行列	13
2.5	アミノ酸の突然変異	15
2.6	アミノ酸置換行列 (PAM250)	16
3.1	活性化エネルギー	18
3.2	典型的な代謝経路	19
3.3	糖の分解とエネルギー	23
3.4	食物から排泄物までの異化の 3 段階	24
3.5	解糖系の経路	25
3.6	クエン酸回路の経路	26
3.7	解糖系とクエン酸回路の中間生成物と、生合成反応系の関係	27
3.8	光合成回路	28
4.1	クエン酸回路の PATHWAY マップ	30
4.2	ENZYME データベースのエントリー例	32
5.1	EC 2.3.1.61 遺伝子 (Dihydrolipoamide S-succinyltransferase) の系統樹 . . .	38
5.2	EC 1.2.4.2 遺伝子 (Oxoglutarate dehydrogenase (lipoamide)) の系統樹 . . .	38
5.3	EC 1.8.1.4 遺伝子 (Dihydrolipoamide dehydrogenase) の系統樹	39
5.4	系統樹内の全ブートストラップ値の中で、95%以上である確率	41
5.5	EC 1.1.1.37 遺伝子 (Malate dehydrogenase) の系統樹	42
5.6	EC 4.1.3.6 遺伝子 (Citrate lyase) の系統樹	42
5.7	EC 4.2.1.3 遺伝子 (Aconitate hydratase) の系統樹	43
5.8	EC 1.1.1.42 遺伝子 (Isocitrate dehydrogenase (NADP+)) の系統樹	43

5.9	EC 1.2.7.3 遺伝子 (2-Oxoglutarate synthase) の系統樹	44
5.10	EC 6.2.1.5 遺伝子 (Succinate-CoA ligase (ADP-forming)) の系統樹	44
5.11	EC 1.3.99.1 遺伝子 (Succinate dehydrogenase) の系統樹	45
5.12	EC 4.2.1.2 遺伝子 (Fumarate hydratase) の系統樹	45
5.13	EC 6.4.1.1 遺伝子 (Pyruvate carboxylase) の系統樹	47
5.14	EC 4.1.1.49 遺伝子 (Phosphoenolpyruvate carboxykinase (ATP)) の系統樹	47
5.15	EC 4.1.1.32 遺伝子 (Phosphoenolpyruvate carboxykinase (GTP)) の系統樹	48
5.16	EC 4.1.3.7 遺伝子 (Citrate (si)-synthase) の系統樹	48
5.17	EC 6.2.1.4 遺伝子 (Succinate-CoA ligase (GDP-forming)) の系統樹	49
5.18	4 分類のクエン酸回路の中での位置	51
6.1	生物種中に存在するアミノ酸由来遺伝子と、光合成由来遺伝子の数	52
6.2	Triose-phosphate isomerase 遺伝子の系統樹	53
6.3	EC1.1.1.37 遺伝子 (Malate dehydrogenase) の系統樹	54
6.4	EC6.2.1.5 遺伝子 (Succinate-CoA ligase (ADP-forming)) の系統樹	54
6.5	EC4.2.1.2 遺伝子 (Fumarate hydratase) の系統樹	55
6.6	EC1.8.1.14 遺伝子 (Dihydrolipoamide dehydrogenase) の系統樹	56
6.7	EC2.3.1.61 遺伝子 (Dihydrolipoamide S-succinyltransferase) の系統樹	57
6.8	EC1.2.4.2 遺伝子 (Oxoglutarate dehydrogenase (lipoamide)) の系統樹	57
6.9	EC2.3.1.61 遺伝子 (アミノ酸生合成由来) のアライメント	59
6.10	EC4.2.1.3 遺伝子 (光合成回路由来) のアライメント	60

表 目 次

1.1	1999 年までに全塩基配列が決定された生物	2
4.1	EC 番号の最上階層 (最左桁) の分類単位	29
4.2	pathway フィールド切出の例	33
4.3	分類結果	35
5.1	アミノ酸生成酵素由来と思われる遺伝子の数	37
5.2	光合成回路由来で系統樹作成に利用した数	40
5.3	炭水化物分解系由来と推測された遺伝子の数	46
5.4	クエン酸回路固有と推定された遺伝子数	50

第 1 章

序章

1.1 背景

近年、ヒトゲノム計画 [17] の進展に代表されるように、多種多様な生物の塩基配列が決定され、データ量は指数的に増加し続けている [6]。本年内にもヒトゲノムの全塩基配列が決定される見通しが立っている。

一方、ヒト以外の生物種では、1995 年にインフルエンザとマイコプラズマの全塩基配列決定に始まり、1999 年までに 25 種類の生物種の全塩基配列が決定されている [3]。26 種類の生物種の内訳は、真核生物、原核生物、古細菌の 3 界にまたがり、12 カテゴリーに渡っている (図 1.1 参照)。

このように多くの生物種の全塩基配列が解明されたことにより、生物種間の全塩基配列同士を比較出来るようになった。これにより異なる生物種の遺伝子同士の対応関係が明らかになり、遺伝子の由来を推定出来るようになった。このように全ゲノムを比較することで、遺伝子の由来を推定する「比較ゲノム解析」が近年発展してきた [1][2]。また、比較ゲノム解析によって遺伝子の機能を推測するような研究も行われ、成果が上がっている。

1.2 目的

従来、生物種間で共有する遺伝子間の差異を比較することにより、進化系統関係を定量化するよう試みられてきた。しかし、単遺伝子の系統と遺伝子の総体である生物種の系統とは一致しないことが知られている [12] [13]。この原因として、トランスポゾンなどの由来が異なる遺伝子が混在することや、遺伝子重複により、途中から違う機能で使われてきたことなどが挙げられる [9]。

category	species	publication
Higher plants	<i>Arabidopsis thaliana</i>	1999
Protozoan	<i>Plasmodium falciparum</i>	1999
Nematode	<i>Caenorhabditis elegans</i>	1998
Fungi	<i>Saccharomyces cerevisiae</i>	1997
Proteobacteria	<i>Escherichia coli</i> K-12	1997
	<i>Haemophilus influenzae</i> Rd	1995
	<i>Helicobacter pylori</i> 26695	1997
	<i>Helicobacter pylori</i> J99	1999
	<i>Rickettsia prowazekii</i>	1998
Gram-positive bacteria	<i>Bacillus subtilis</i>	1997
	<i>Mycoplasma genitalium</i>	1995
	<i>Mycoplasma pneumoniae</i>	1997
	<i>Mycobacterium tuberculosis</i> H37Rv	1998
Chlamydia	<i>Chlamydia trachomatis</i>	1998
	<i>Chlamydia pneumoniae</i>	1999
Spirochete	<i>Borrelia burgdorferi</i> B31	1997
	<i>Treponema pallidum</i>	1998
Cyanobacteria	<i>Synechocystis</i> PCC6803	1996
Radioresistant bacteria	<i>Deinococcus radiodurans</i> R1	1999
Hyperthermophilic bacteria	<i>Aquifex aeolicus</i>	1998
	<i>Thermotoga maritima</i> MSB8	1999
Archaea	<i>Methanococcus jannaschii</i>	1996
	<i>Methanobacterium thermoautotrophicum</i> deltaH	1997
	<i>Archaeoglobus fulgidus</i>	1997
	<i>Pyrococcus horikoshii</i> OT3	1998
	<i>Pyrococcus abyssi</i>	1999
	<i>Aeropyrum pernix</i> K1	1999

表 1.1: 1999 年までに全塩基配列が決定された生物
[3] より引用

当研究では、この問題を解決するために、一緒に進化してきたと思われる遺伝子群を既存データベースに登録されている情報を利用して分類する方法論を提起し、検証を試みた。その際、生物の基本的な遺伝子セットであり機能の単位である代謝系に注目し、代謝系遺伝子の由来を分類し、分類群内での系統について考察を行った。

本論文の構成は以下の通りである。第2章では、これまでに確立されてきた分子系統学の理論を述べる。第3章では、生物の代謝系、特にクエン酸回路に関係のある説明を行う。第4章では、遺伝子データベースから得られた情報について解説した。第5章では、4章から得られた情報を元に分子進化系統樹を作成し、検証を行った。第6章では、5章の結果から特徴的な部分に注目し考察した。最後に、第7章でまとめた。

第 2 章

分子系統学における前提と諸問題

2.1 分子進化

2.1.1 進化概説

かつて、生物種の分類を考える際には、形態の類似度によって行われていた。[15]

しかし、魚類と鯨類は由来が異なっているにも関わらず、水中遊泳を行うという生活形式を共有しているため、極めて類似した形態に進化してきた。このように形態から分類を行うことには、判断する者の主観に左右されやすく、定量化することが難しいという欠点を持っている。このように形態で分類するという方法は、近代分類学の祖である Carl von Linne に始まり、現在に至るまで続いている。[15] この方法が生まれた背景として、全ての生物は下等な生物が時間と共に進歩し、高等な生物に至ったという思想がある。つまり、進歩するに従ってより複雑な形態を獲得していくため、形態が似ている生物種同士は近縁であるという考え方である。この考え方は、高等生物は太古の昔に誕生し進歩しつづけてきており、下等生物は、最近誕生したものであるという事を意味している。これに対し、Charles Darwin は、彼の著書「種の起源」において、現在地球上に生息する全ての生物は、1つ、もしくは、少数の祖先から進化してきたと述べた。その後、生物の遺伝を司る媒体は、DNA、あるいは RNA といった核酸であることが解明され、分子生物学が発展すると共に彼の説の信憑性が高まり続けている。そして現在、多くの生物学者は、生物種は共通の祖先から進化して来たという前提の元に進化に関する研究を行っている。[13]

2.1.2 分子時計

共通の祖先から種が分岐してきたならば、分岐後に生物種の違いをおこしてきたのは何であろうか。

それは、遺伝媒体である核酸が突然変異を起すことによって、種が分岐してきたと考えられている。従って、生物種の違いは核酸の塩基配列の違いに近似させることができる。そして、1962年に Emile Zuckerkandl と Linus Pauling によって、突然変異に関する重要な性質が発見された。一般には「分子時計」という名で知られている現象である。[9] 彼らは、アミノ酸配列の置換数と、化石から得られた分岐時期に関して図 2.1.2 の関係を導き出した。現在の知見からこのグラフに関して言えることは、塩基配列から翻訳されるアミノ酸配列の突然変異は、時間に比例して蓄積しているということである。これは、分子置換速度は一定であると言い換える事が出来る。

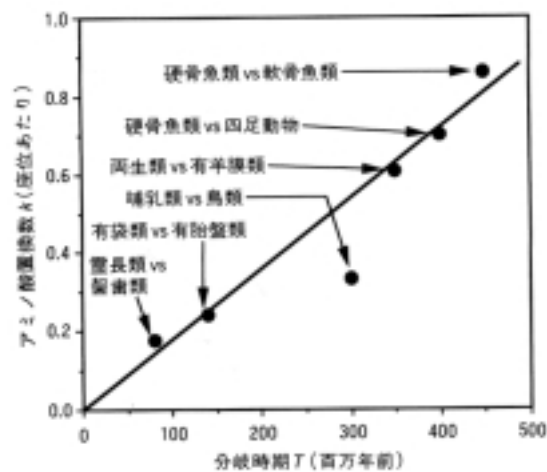


図 2.1: ヘモグロビンの分子時計
(文献 [12]1.4 より)

この性質は、同じ機能のタンパク質を持つ 2 種類の生物種のアミノ酸配列間の置換数と、そのタンパク質の分子置換速度が分かれば、その遺伝子が分岐した時期が推定できる事を示す。また、分子置換速度が分からない場合でも、複数種のアミノ酸配列の置換数が分かれば、相対的な分岐位置を推定でき、分類できる。

この原理の元に、Fitch と Margoliash は酵母からヒトに至るまでの多様な生物種のチトクローム c 分子の分岐位置を推定した結果、その当時考えられていた系統関係と一致していたため、方法としての信頼性が認められるようになった。[12]

2.1.3 自然淘汰説と分子進化の中立説

Darwin が自然淘汰説を提唱して以来、進化は自然淘汰によって行われていると考えられてきた。自然淘汰説では、突然変異を 2 つの種類に分けて考えている。

第一は、変異がおきることによって機能が向上するような変異である。このような変異は他の個体に比べて生存する可能性が高くなり、子孫を残す可能性が上がるため、時間が経過するに従って子孫の数が増加してくる。第二は、変異が起きることによって機能が低下、もしくは、有害な影響を起す変異である。このような変異は、死亡する確率が高く、子孫を残さない可能性が高くなり、次の世代に受け継がれない可能性が高い。このような理由により、生活する環境に有害である変異は個体の死亡と共に除去され、有利な変異は子孫が増加する事によって増加し、時間の経過と共に有利な形質のみが生物集団中に固定され、生物は進化してきたと考えている。

これに対し、木村資生は多くのタンパク質の置換に関するデータを基に、1968 年、分子進化の中立説を提唱した。[10] 中立説では、分子の突然変異に限り、3 つの種類に分けて考えている。

自然淘汰説で述べた、有利な突然変異、不利な突然変異の 2 つと、生存に有利にも不利にもならない中立な突然変異である。生存に不利な突然変異は自然淘汰説で述べられている理由により除去される。一方、生存に有利な突然変異は、生存に有利ではあるが、生物集団中では極めて少数であり、集団の中で固定される確率は無視できるほど低い。集団中で大多数をしめる突然変異は、中立な突然変異であり、その突然変異が固集団注で固定されるのは遺伝子の機能にはよらず、偶発的な事象であると論じた。

なお、ここで注意すべき点としては、中立説が述べているのは単体の分子の突然変異に関してであり、表現型や、抽象化された形質に関しては自然淘汰説とは矛盾しない事があげられる。

2.1.4 中立突然変異と進化速度

中立説の論拠は進化速度に基づいて述べられているため、始めに進化速度に関する式を以下に記す。アミノ酸の座位辺り起きている置換数を k 、分岐時期を T 、進化速度を v としたときに、

$$k = 2vT \quad (2.1)$$

$$v = \frac{k}{2T} \quad (2.2)$$

とすることができる (図 2.1.2 参照)。係数 2 は、2 つのアミノ酸配列が分岐後に、それぞれ独立に置換した和をとっている。

二倍体の染色体をもつ生物が N いる集団において、中立な突然変異が集団に固定される確率は $\frac{1}{2N}$ である。中立突然変異が起きる頻度が μ とした時、集団内で起きる突然変異の総数は $2N \cdot \mu$ となり、集団内に固定される中立突然変異、すなわち、中立突然変異進化速度 v_n は、

$$v_n = 2N \cdot \mu \cdot \frac{1}{2N} = \mu \quad (2.3)$$

となり、中立突然変異が起る確率と集団中に固定される確率が等しくなる。

これらの中立突然変異が、全ての突然変異 μ_T の中で起る割合を f とすると、

$$v = f \cdot \mu_T \quad (2.4)$$

が導き出され、進化速度は、突然変異発生率の増減によって左右される事を意味している。

2.1.5 遺伝子の多様性

式 2.4では、分子進化速度は全突然変異中の中立進化の頻度 f と、突然変異発生率 μ_T によって決定する事を意味している。[12][14]

突然変異発生率について前説で述べたので、当節では、中立進化の頻度に関して論ずることから始める。分子進化の中立説に従えば、有利な突然変異は無視できる程少数であるため、残りの $1 - f$ が不利な突然変異発生率となる。

生体内で機能しているタンパク質は、他のタンパク質と相互作用をする事によってその機能を発揮している。従って、他のタンパク質で代替出来ない重要な遺伝子は、機能が変化したときに生存に不利になる可能性が高い。従って、重要な機能を持った遺伝子である程、不利な突然変異発生率 $1 - f$ が増加するため、分子の進化速度は遅くなる。このように、分子の機能によって、進化速度が減少する事を、機能的制約と呼ぶ。このように機能遺伝子の分子進化速度を減少させて保守的になり、中立遺伝子のみが固定される過程には、新しい機能遺伝子が生じる機構が存在していない。このような問題を解決する機構として、遺伝重複が挙げられる。[12]

生物内には、同じ機能をもったタンパク質が多量に存在しているが、これは、元の遺伝子がコピーされる事によって増殖してきた。同じ機能を持ったタンパク質が存在していれば、その内の一つが突然変異によって機能しなくなっても生存に不利にはならない。従って、中立突然変異頻度 f が 1 となり、進化速度は速くなり、新たな機能をもった遺伝子が生じる可能性が生じてくる。このように、遺伝子がコピーを増やす事を遺伝子重複といい、必要な機能は保存しつつ、新しい機能遺伝子を作るものと考えられている。

2.1.6 中立進化の検証

これまで、中立進化の理論に沿った説明をしてきたが、理論提唱以来多くの反論を受けてきた。反論の多くは、分子進化についても、自然淘汰が働いているという主張である。[10] 前節で述べた遺伝子重複は、コピーをとることで、進化速度が速くなったという主張である。この場合、その遺伝子の機能を突然変異によって失っても、生存には中立であるため、進化速度が速くなるという主張であるが、機能がなくなった遺伝子はどうなるのだろうか。

分子進化に自然淘汰説を適用すれば、機能がなくなった遺伝子は、正にも負にも淘汰は働かない。したがって、進化速度は最小になり、集団に定着される確率が少なくなり、中立進化説の結論とは正反対になる。

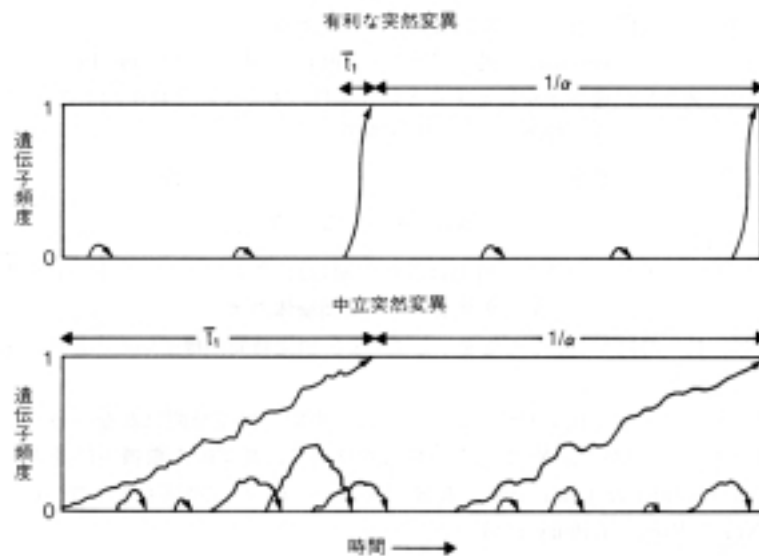


図 2.2: 中立進化と、適応進化
(文献 [14] より)

分子進化の中立説の発表以来、様々な配列データに関する情報が検討されてきた。その結果、多くの生物種で機能のある遺伝子に配列が似ているが、現在は機能を持っていない配列が大量に発見された。この様に、遺伝子重複によってコピーされた後、突然変異によって機能を失なった遺伝子を、偽遺伝子という。この偽遺伝子の進化速度と、機能遺伝子の進化速度を比較したところ、偽遺伝子の方が進化速度が速くなるという結果が出され、中立説の正当性が認められ、現在多くの研究者に支持されるに至った。

2.1.7 機能遺伝子の分子進化と系統

前節では、重複した遺伝子が機能を失った場合についての議論であったが、本節では、新たな機能を獲得した場合について論じる。

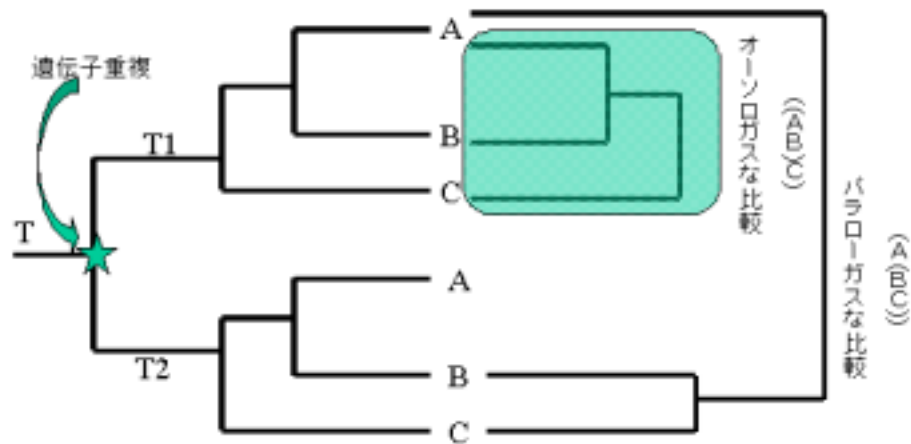


図 2.3: 機能遺伝子の進化

図 2.3は、現在生存している生物 A, B, C に、 $T1, T2$ という遺伝子がある時、その共通祖先遺伝子 T が遺伝子重複を起して別々の機能になったあと、種分岐が起きた事を仮定している。重複によって別々の機能が生じた時点で、 $T1, T2$ にかかる機能制約が異なってくるため進化速度はそれぞれ異なってくる。しかし、 $T1, T2$ 共に共通の祖先遺伝子 T から発展してきたため、配列比較をすると高い類似性を示してしまう。そのため、我々が A, B, C の系統を調べようとした時、 $T1$ 遺伝子と $T2$ 遺伝子を混在して比較してしまい、系統関係が誤って推定されてしまう危険性がある。このように、機能制約、及び、進化速度の異なる遺伝子を比較してしまう事を、paralogous な比較という。それに対して、同じ機能を持ち、機能制約が等しい遺伝子同士を比較する事を、orthologous な比較という。多くの分子系統学の研究結果では、遺伝子の系統関係が我々が一般的に種と呼んでいるものの系統関係を反映していないとの報告が多く出ている。この原因としては、paralogous な比較をしているため、誤った結果が導き出されているという事が多い。[13] この場合は、種の系統と遺伝子の系統が異なっているのではなく、比較方法が誤っていると考えられる。これに対し、種の系統と遺伝子の系統が異なるという現象は別の理由により生じている。

第一に、遺伝子の分化が起きて世代が経った後に種分化が起きる。この現象自体は分岐

時期が異なっているに過ぎず、分化の順序¹にずれは生じないが、遺伝子分岐の直前に種分岐が起きた場合に、分岐の順序が異なってしまう。

第二に、遺伝子が異種の個体にコピーされてしまう場合がある。このような水平移動が起きた遺伝子を比較しても、種の系統関係は一切反映されない。

第三に細胞内共生があげられる。真核生物内に存在するミトコンドリアや葉緑体は、元来別の生物が細胞内に取り込まれることで細胞内器官となったと考えられている。これらの細胞内器官は独自に遺伝子を持っており、その一部は核に取り込まれている。このため、共生前に存在した遺伝子と、共生後に取り込まれた遺伝子間の系統樹のトポロジーは異なってしまう。系統樹に関して考察するときにはこのような点に考慮しなくてはならない。

¹系統樹のトポロジーの事

2.2 配列比較の方法

2.2.1 アラインメントの概念

これまで、分子進化の概念について論じてきた。本章以降では、配列を比較する方法論について例を挙げながら論じていく。Mycoplasma pneumoniae（以下、マイコプラズマ）と、Mycobacterium tuberculosis（以下、結核菌）は共にグラム陽性細菌である。この両者は、tRNA を切断する遺伝子 Ribonuclease を共に持っている。この両遺伝子のアミノ酸配列の長さは、マイコプラズマが 118 なのに対し、結核菌は 125 である。この両者の長さの差は、突然変異によってアミノ酸が挿入したか欠失したかのいずれである。（以後、挿入もしくは欠失の事をギャップと呼ぶ）このように長さが異なると、どこで置換があったのか比べる事ができないため、位置あわせをしなければならない。以下に、マイコプラズマのタンパク質 50 残基目からと、結核菌のタンパク質 58 残基目以降を示し、位置あわせについて説明する。

先頭から並べると

50

KTKYKLAVER マイコプラズマ

@

KSVGSAVER 結核菌

58

のようになり、先頭の一文字目のみが一致していてそれ以外はすべて異なり、置換率は $\frac{9}{10}$ と非常に高い。この両者をよく見比べてみると、終端から 4 文字の AVER は共に一致しているため

50

KTKYKLAVER マイコプラズマ

@ @@@@

KS-VGSAVER 結核菌

58

と並べ直し、途中にギャップを挿入すると、多くのアミノ酸を一致させる事が出来る。この場合²置換率は $\frac{5}{10}$ という結果になる。

²文字の一致という観点からすると、2 文字目から 6 文字目まで

のどの位置にギャップが入ってもいいことになるが、アミノ酸の性質の類似度によって上記の位置に決定した。後節で詳細を述べる。

この様にギャップを挿入することによって、置換率が少なくなるように位置を合わせることをアラインメントと呼ぶ。

2.2.2 アラインメントのアルゴリズム

前節の例では、アラインメントを手作業で直観的に行った。本節では、機械的に処理出来るアルゴリズムと、理論的背景について論じる。

2つの配列のアライメントをとる時、共通の祖先から最小の置換数で現在の配列に至った考えることができる。³ 従って、2つのアライメントは出来るだけ文字列が一致し、ギャップの数が少なくなるような並べ方になる。2つの配列のアライメントを任意に取った時、生じたギャップの数を g 、ギャップが入ることの不当性を α 、文字列の不一致数を d 、文字列が不一致であることの不当性を β 、その任意のアライメントの不当性を P と定義した時

$$P = \alpha \cdot g + \beta \cdot d \quad (2.5)$$

となり、全ての並べ方の中で P が最小となるようなアライメントが最適である。Needleman と Wunsch は、この性質を動的計画法 (Dynamic Programming) に適用したアルゴリズムを提唱した。[14][18]

このアルゴリズムでは、始めに2本の配列を2次元行列の軸とする。(図2.4参照)

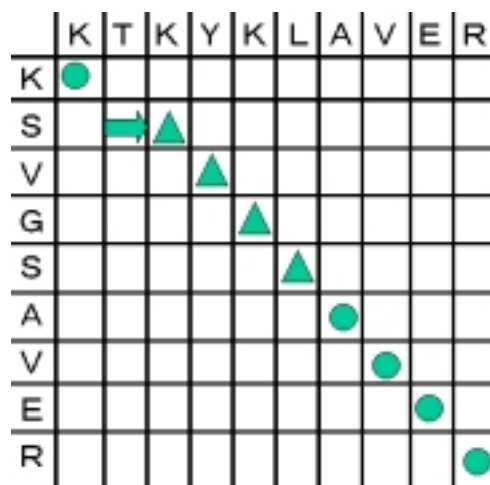


図 2.4: アラインメント行列

³これを最小進化の考え方と呼ぶ。[16]

図中で丸印の所が文字列一致、三角印の所が文字列不一致、矢印の所が、ギャップを表している。この図が意味している事は、行列の始点である 1 行 1 列地点から、終点である 9 行 10 列地点までの全ての経路の中で式 2.4 が最小となる経路を求めるという組み合わせ最適化問題に帰着する点である。これを以下のように定式化する。

A 配列と B 配列の 2 本の配列の長さをそれぞれ、 m, n 、任意の配列要素を i, j 、従って、配列 A を $a_1, \dots, a_i, \dots, a_m$ 、配列 B を $b_1, \dots, b_i, \dots, b_m$ と定義し、変数 a_i, b_j には、それぞれ、アライメントの不当性 P_i, P_j をセットする。不一致の起りにくさを $\sigma(a_i, b_j)\beta$ を表し、a と b が等しければ β 、不一致ならば 0 を取るような関数を定義する。この時、アライメントの良さ P_{ij} は、以下の式で表される。

$$\text{ギャップでない場合} \dots P_{ij} = \min\{P_{i-1} + \sigma(a_i, b_j)\beta\} \quad (2.6)$$

$$A \text{ にギャップが入る場合} \dots P_{ij} = \min\{P_{i-1,j}\alpha\} \quad (2.7)$$

$$B \text{ にギャップが入る場合} \dots P_{ij} = \min\{P_{i,j-1}\alpha\} \quad (2.8)$$

ここまで述べてきたのは、2 配列間の比較であったが、遺伝子の系統関係を調べる際には、3 配列以上の比較を行わなければならない。3 配列以上の比較のアライメントをマルチプルアライメント呼ぶ。マルチプルアライメントでは、2 次行列を多次元空間に拡張し同様に経路を求めて行く。組み合わせ最適化問題であるという点においては、変わりがないが、多次元になると経路数が爆発的に増加するため、系統樹を作るという目的では適当ではない。アライメントを取る際、文字列の一致が連続する部分は、共に同じ祖先配列である可能性が高いという性質を利用し、あらかじめ一致する文字列が続いている部分を配列内から抽出しておく。このような部分配列⁴周辺領域のみに限定して、動的計画法を実施することによって、計算量を大幅に減少させ⁵、マルチプルアライメント [18] の近似解を求める方法を行っている。

⁴このような部分配列をダイアゴナルと呼ぶ

⁵あらかじめ、全領域に部分配列のような目次をつけ高速化する手法を indexing 法という。

2.2.3 アミノ酸の突然変異

前節では、ギャップとアミノ酸文字列の不一致へのペナルティである指標である、 α と β を仮定していた。当節では、アミノ酸の不一致に関しての考察、及び、突然変異のモデル化について論じる。タンパク質が細胞内で機能する時、タンパク質の立体構造が機能を決定している。タンパク質の立体構造は、要素であるアミノ酸のつながり方によって決まるため、アミノ酸の大きさ、極性、水和度などの、物理的・化学的性質が重要になっている (図 2.5 参照)。[21]

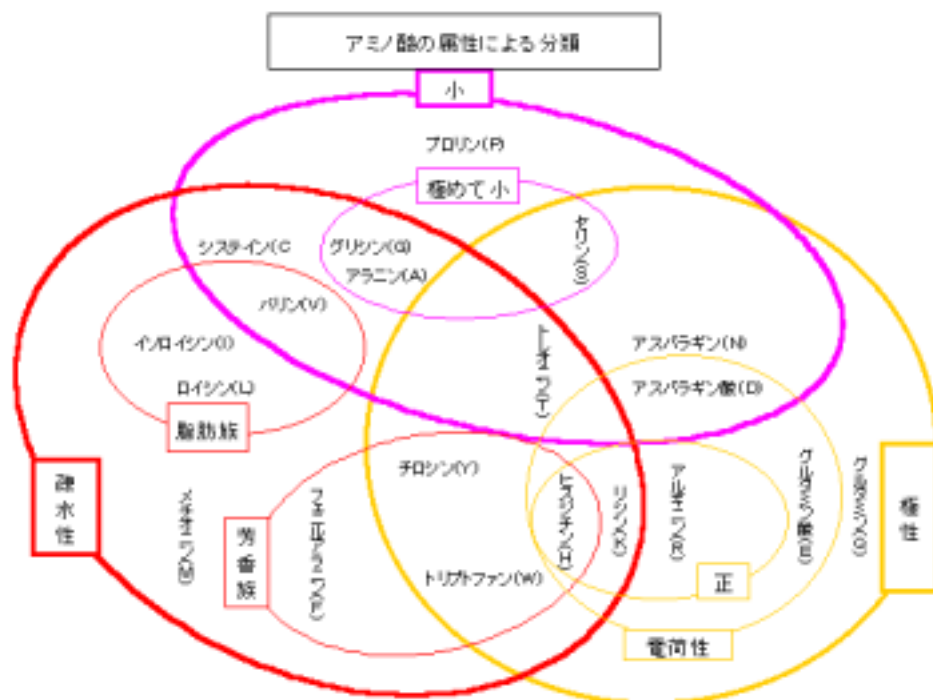


図 2.5: アミノ酸の突然変異

例を挙げると、アミノ酸が突然変異で置換されたとき、大きさや極性などが元のものと大きく異なるとアミノ酸同士のつながり方の角度などがずれてしまい、立体構造が変化し、従来の機能が消失してしまい負の淘汰を受けてしまう。一方、類似したアミノ酸に置換した時には、立体構造への影響が少なく、中立な突然変異である可能性が高い。

このような理由により、置換のペナルティーは一律ではなく、置換の種類によって変えるべきである。このような観点から、Dayhoff は、多数のタンパク質のアミノ酸配列から

第 3 章

代謝経路の構造及び進化について

3.1 代謝

生物が生存していく上で最も重要な事として、細胞本体を構成するのに必要な分子を作り出す事と、生体活動を行うのに必要なエネルギーを生産する事がある。両者共、自然界から取り込んだ物質を化学反応させることで実現している。化学反応には2種類あり、分子をさらに小さい分子に分解する異化と、分子同士を結合して新しい分子を作り出す同化（生合成）である。異化では、分子は分解の際、分子同士の結合エネルギーを放出する。この放出された結合エネルギーを ATP 等のエネルギー運搬物質に蓄える役割を果たし、さらに、有用な小分子を作り出す機能も担っている。同化では、異化によって作り出されたエネルギーを利用して、細胞の要素となる分子を合成する。異化と同化の反応を合わせて代謝と呼び、生物が生存していく上での秩序を提供している。これらの生体内反応は、酵素と呼ばれるタンパク質によって制御されている。

一般に異化はエネルギーを生産する反応であるが、反応初期に反応を行うためのエネルギーを消費する（図 3.1 参照）。酵素は、この活性化エネルギーの値を小さくする触媒の役目を果たす。この作用により、化学反応が始まり、初期の活性エネルギーよりも大きなエネルギーを得ることができる。このような代謝反応は連続しておきる事が多く、図 3.1 ではその一連の経路として描いている。

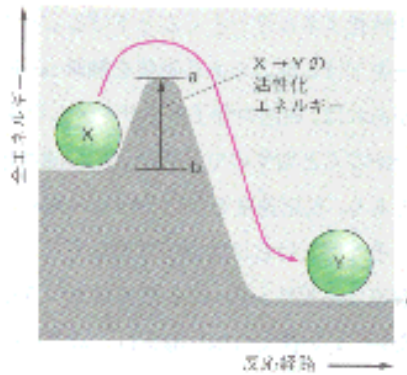


図 3.1: 活性化エネルギー
(文献 [22] より)

3.2 栄養物質の分解

3.1節で、エネルギーは食物分子を分解する事によって得られた結合エネルギーを ATP 等のエネルギー運搬分子として蓄えると述べた。代謝反応の原料となる食物分子として主要なものには、糖、タンパク質、脂肪がある。これらの中で、エネルギーを生産する上で最も重要な食物分子は、糖である。糖は代謝経路に従って段階的にエネルギーを生産し、最終的には二酸化炭素と水になる（図 3.2 参照）。

糖を分解してエネルギーに変換する過程は、3つの段階に分類される（図 3.2 参照）。

まず始めに、多糖を単糖に分解して、細胞内で汎用的に利用できる形にする。第二段階として、グルコース分子をピルビン酸分子2個に分解することによって、エネルギー運搬物質である ATP と NADH を生産する解糖系という一連の反応群を形成している。解糖系は、反応の際酸素を必要としないため、嫌気性細菌を含んだほとんどの生物で普遍に行われている。嫌気性細菌は後述するクエン酸回路と電子伝達系の反応を行わないため、エネルギー生産の全てを解糖系でまかなっている。このことから、生物進化の初期からエネルギー生産の役割を担ってきたものと考えられている。[?] 第三段階で行われる反応は、真核生物において、ミトコンドリア内細胞内で行われている。¹この段階は、さらにクエン酸回路と、電子伝達系の2つに分かれる。これらの経路は、3.3節で詳細に説明を行う。

¹好気性細菌の細胞膜周辺でも同様の反応が認められている

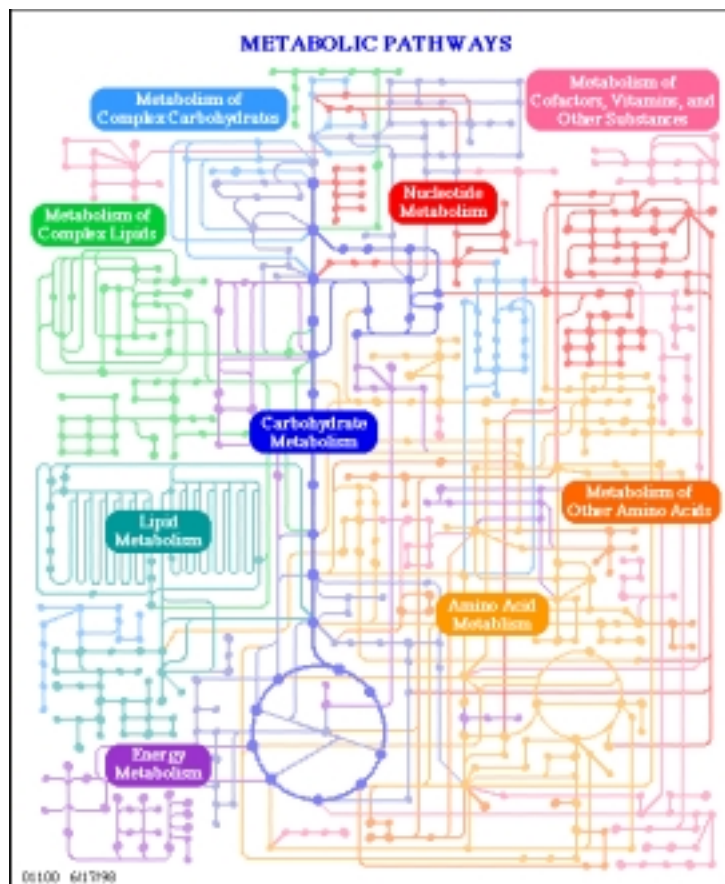


図 3.2: 典型的な代謝経路
(文献 [3] より)

3.3 酸化リン酸化

3.3.1 クエン酸回路

解糖系の最終生成物はピルビン酸であるが、ピルビン酸の時点でも高い化学エネルギーを保持している。クエン酸回路では、このピルビン酸をや、脂肪酸²を酸化することによりエネルギーを得ることができる (図 3.3.1 参照)。

ピルビン酸は、直接酸化されるのではなく、アセチル CoA を経てピルビン酸を作る。このピルビン酸が段階的に酸化を行い、最終的生成物オキサロ酢酸に至るまでにエネルギー運搬物質である NADH₃ 分子、GTP 1 分子を生産する。

そして、オキサロ酢酸と新たなアセチル CoA と反応しピルビン酸が生成されることによって、回路として成立する。解糖系とクエン酸回路は、エネルギー物質として重要なのはこれまで述べてきた通りであるが、中間生成物も他の分子が生合成される際の前駆物質として使用されている (図 3.3.1 参照)。

²脂肪酸は、図 3.2 で記載されている脂肪に由来する

3.3.2 エネルギー運搬物質

[23] エネルギー運搬物質が、他分子の生合成にエネルギーを供給する際には、エネルギー運搬物質の分解と、生合成反応が同時に起こることによってエネルギーの受け渡しが行われ、この反応の事を共役反応と呼ぶ。解糖系で作られたエネルギー運搬物質 ATP(アデノシン 5' 三リン酸) は、核酸の一種であるアデニンと、五単糖のリボースにリン分子が3つついた構造をしている。ATP は加水分解することによってリン分子が一つ外れ、ADP(アデノシン 5' 二リン酸) が生成される。この際放出されるエネルギーが、他の生合成反応を促進させている。クエン酸回路によって生成された NADH(ニコチンアミドアデニンジクレオチド) と FADH_2 も、重要なエネルギー運搬物質である。

NADH と FADH_2 は、水素イオン (H^+) が外れる際に高エネルギー電子を放出し、ATP と同様に他の生合成反応を促す。そして、両者の大きな働きとして、ミトコンドリア内膜に存在する酵素群に分解反応で生じた H^+ と高エネルギー電位を送ることによって、ADP を ATP に復活させる事が出来る。この反応を酸化リン酸化と呼び、一連の経路を電子伝達系と呼ぶ。

これまで述べてきた糖の酸化の過程では、グルコース 1 分子が、解糖系、クエン酸回路、電子伝達系を経由することで、ATP30 分子が生産された事になる。

3.4 光合成回路

3.3節では酸素を利用することにより、ATP を効率的に生産している事を説明してきた。一方、生命誕生の時点では、地球上の酸素濃度は極めて低く、酸化的リン酸化が行なえるだけの酸素濃度が充満したのは、生命誕生から 10 億年の時が経ってからだった。[21] この酸素増加の原因になったのが、生物の光合成反応である。

光合成では、太陽の光エネルギーを利用してみずと二酸化炭素とか、糖と酸素を作り出す。この反応は、クエン酸回路の逆反応であり、反応の中間生成物、反応に介在する酵素も同じである (図 3.4 参照)。

これらの反応は、光合成細菌と、葉緑体をもつ真核生物 (以下、植物) において行われている。これらの反応がおきている部位は、光合成細菌では細胞内、植物では葉緑体内部である。

さて、真核生物では、クエン酸回路はミトコンドリア、植物では葉緑体という小器官で似たような反応がおきているが、これはどのように考えられるのだろうか。

ミトコンドリアと葉緑体は、機能だけでなく、構造上にも類似点がある。類似点の中で興味深い事として、細胞膜が 2 重³であり、核とは別に DNA をもち独自に遺伝子の複製を行っている事があげられる。[21][22] これらの特徴から、真核生物本体と、ミトコンドリア、葉緑体は元々別種の生物であったが、ミトコンドリアと葉緑体の祖先生物が、真核生物の祖先生物内に取り込まれ共生するようになったのだと考えられている。この説を細胞内共生説と呼ぶ。

³小器官の祖先生物が細胞内に貫入される際、自らの細胞膜と寄種生物の細胞膜が合わさって 2 重膜になったと考えられている

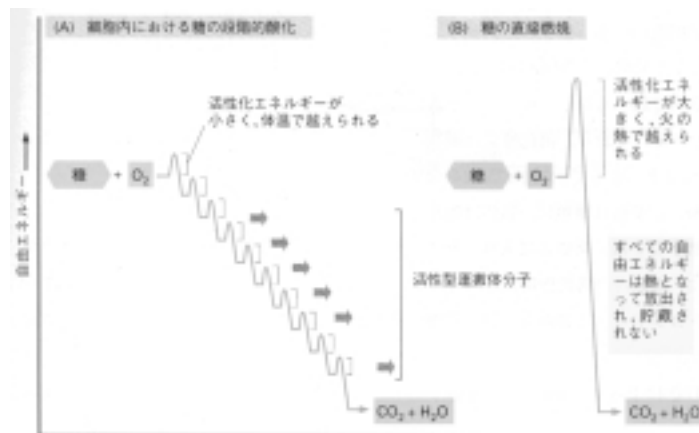
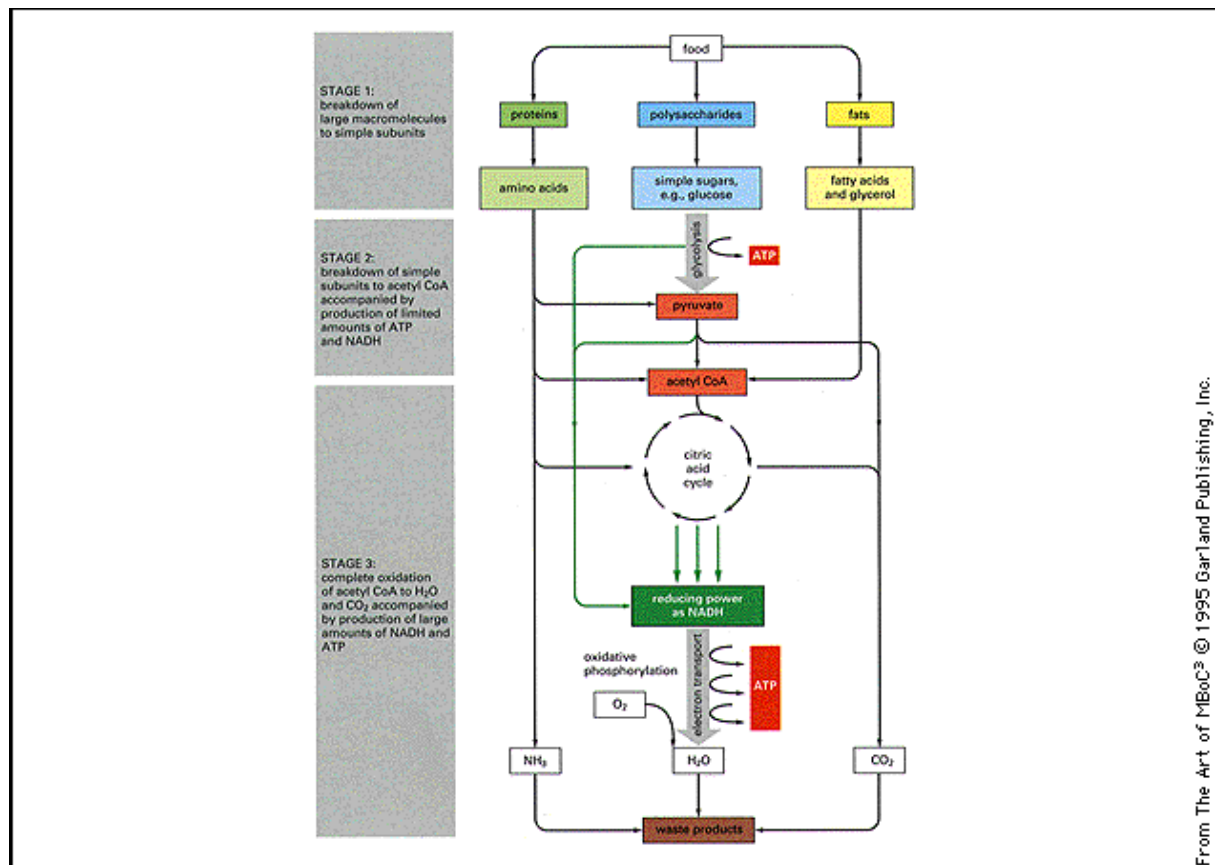
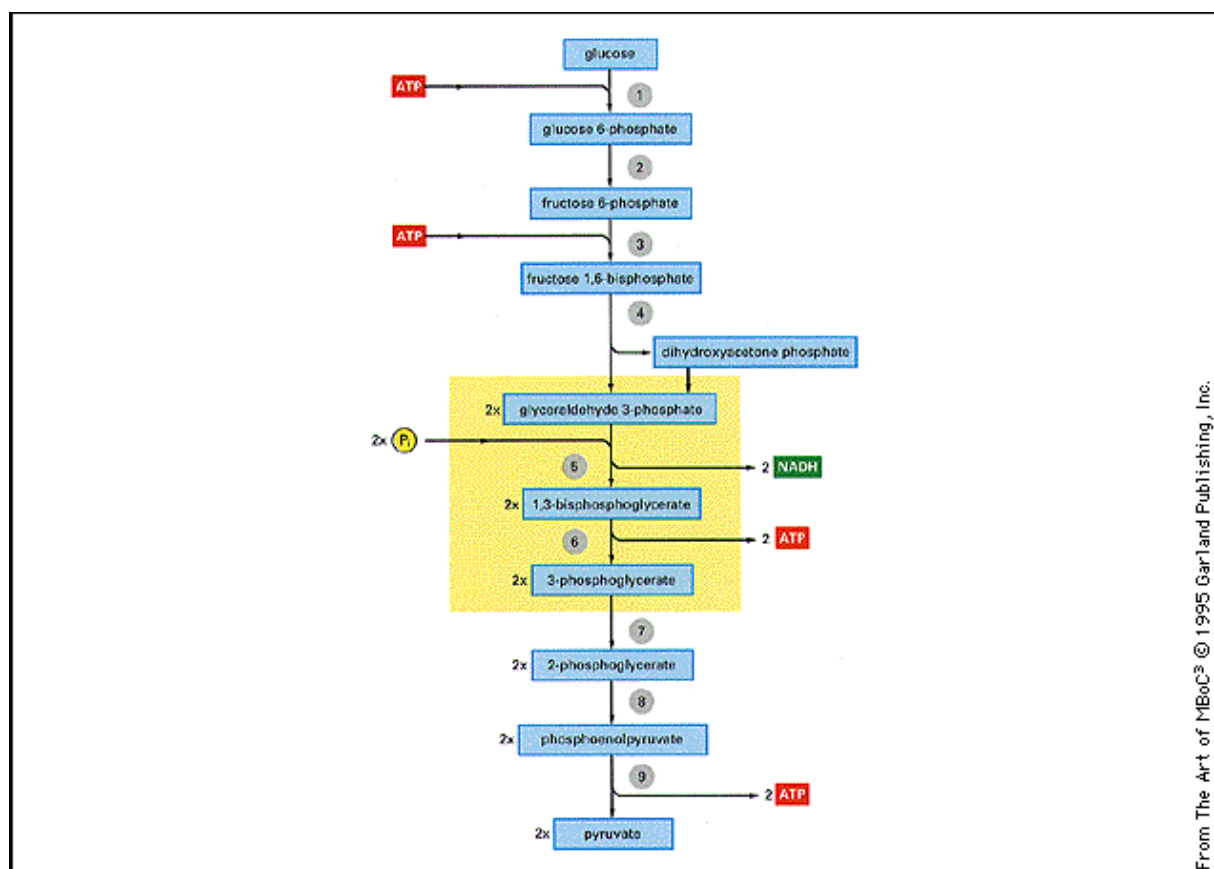


図 3.3: 糖の分解とエネルギー
(文献 [22] 図 4-1 より)



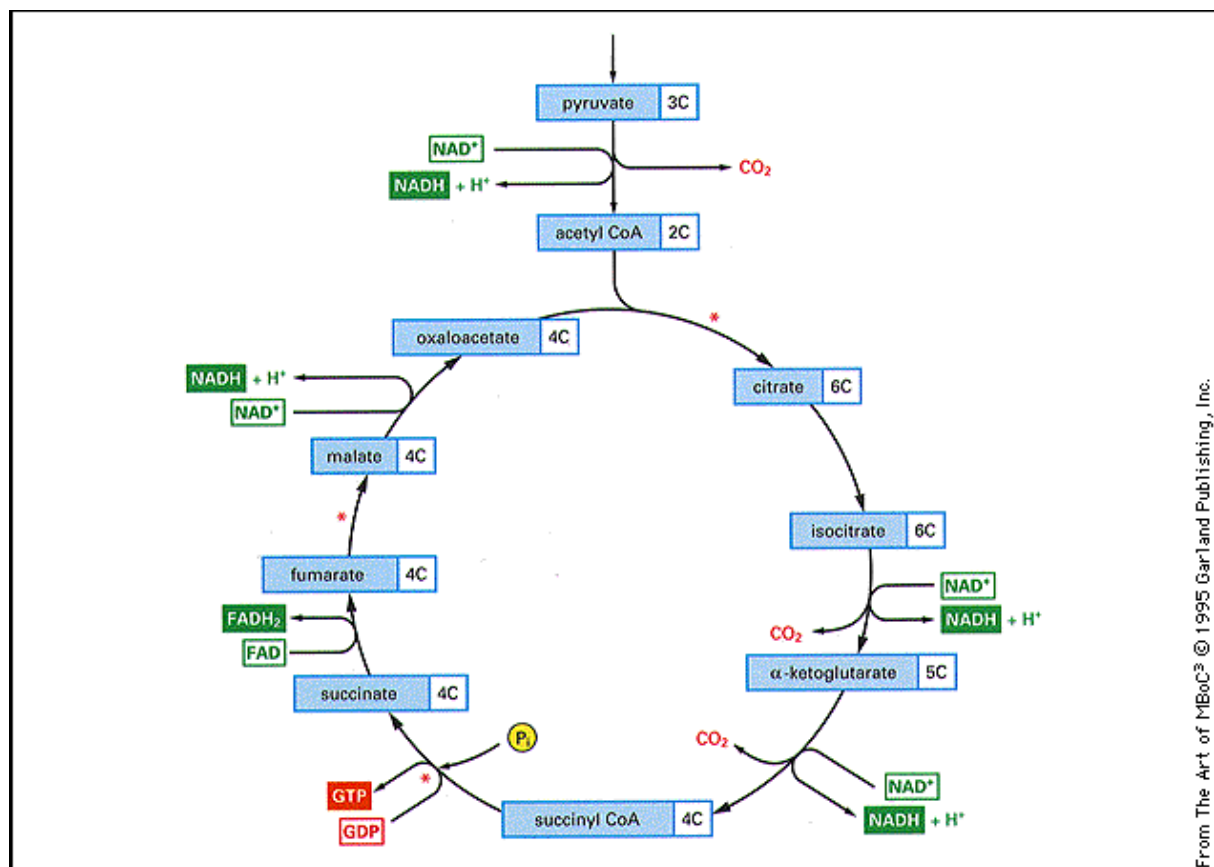
From The Art of MBoC³ © 1995 Garland Publishing, Inc.

図 3.4: 食物から排泄物までの異化の 3 段階
(文献 [21] より)



From The Art of MBoC³ © 1995 Garland Publishing, Inc.

図 3.5: 解糖系の経路
(文献 [21] より)



From The Art of MBoC³ © 1995 Garland Publishing, Inc.

図 3.6: クエン酸回路の経路
(文献 [21] より)

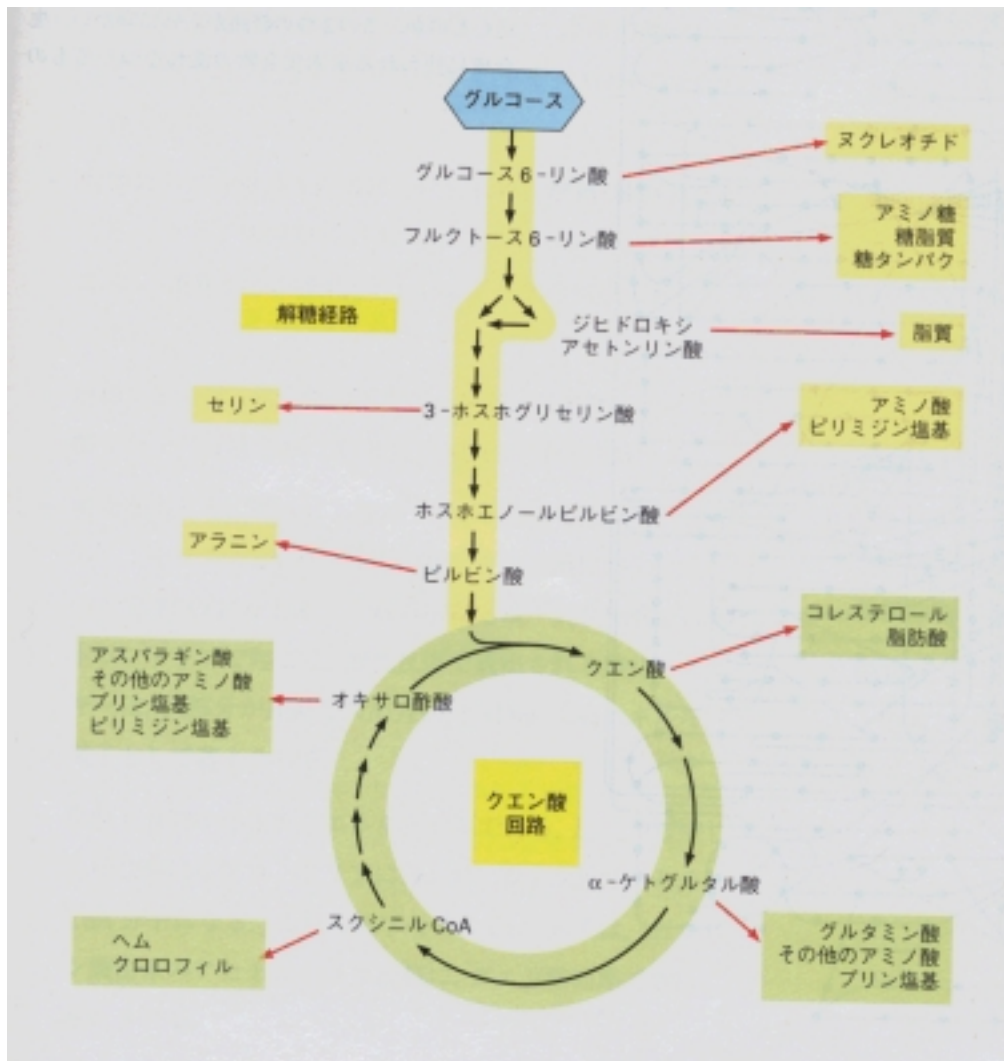


図 3.7: 解糖系とクエン酸回路の中間生成物と、生合成反応系の関係
(文献 [22] より)

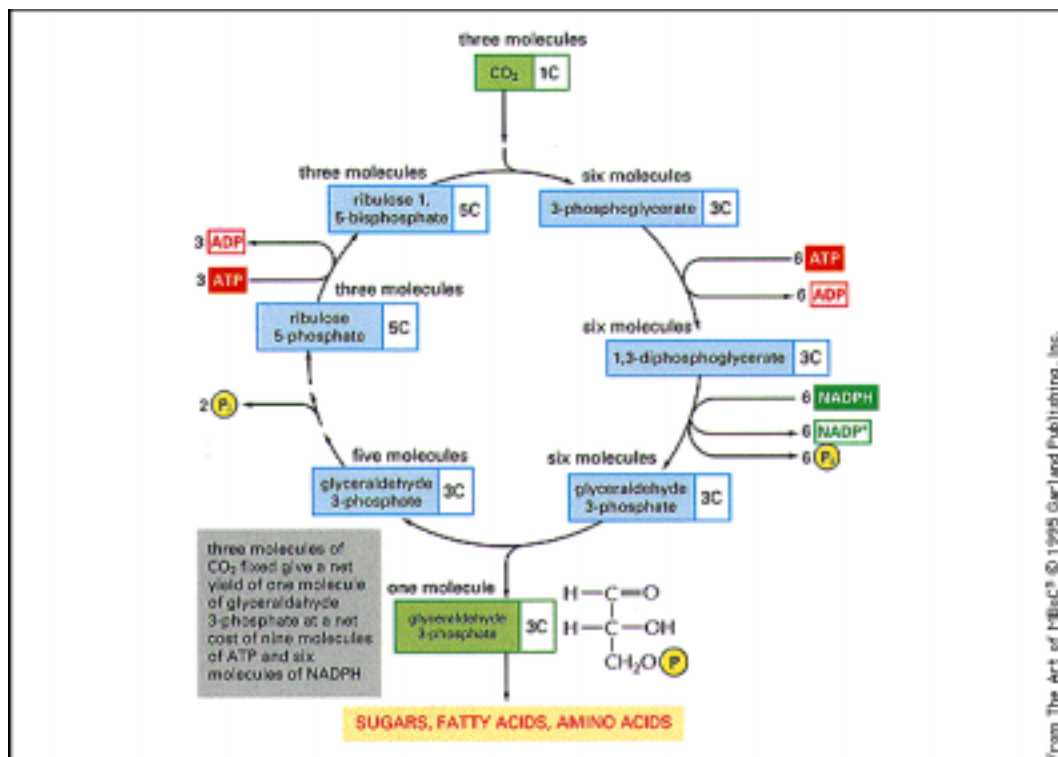


図 3.8: 光合成回路
(文献 [21] より)

第 4 章

既存のデータベースから得られた知見

4.1 データの入手

本研究では、クエン酸回路遺伝子 20 個を対象にした。対象となった遺伝子データは、京都大学化学研究所で構築されている KEGG(Kyoto Encyclopedia of Genes and Genomes)[3] 内の、PATHWAY データベース、及び、ENZYME データベース内登録データを利用した。PATHWAY データベースは、代謝経路をグラフィカルに表現したもので、酵素反応と反応基質へのリンクが張られている (図 4.1参照)。

ENZYME データベースは、酵素反応を元に遺伝子を分類したものであり、図 4.1内で四角枠に囲まれている 4 つ組の数字からリンクが張られている。この 4 つ組の数字は、EC ナンバーと呼ばれ、酵素反応の性質によって、左の桁から階層的に分類されている。[?](表 4.1 参照)

	function
1	Oxidoreductases
2	Transferases
3	Hydrolases
4	Lyases
5	Isomerases
6	Ligases

表 4.1: EC 番号の最上階層 (最左桁) の分類単位

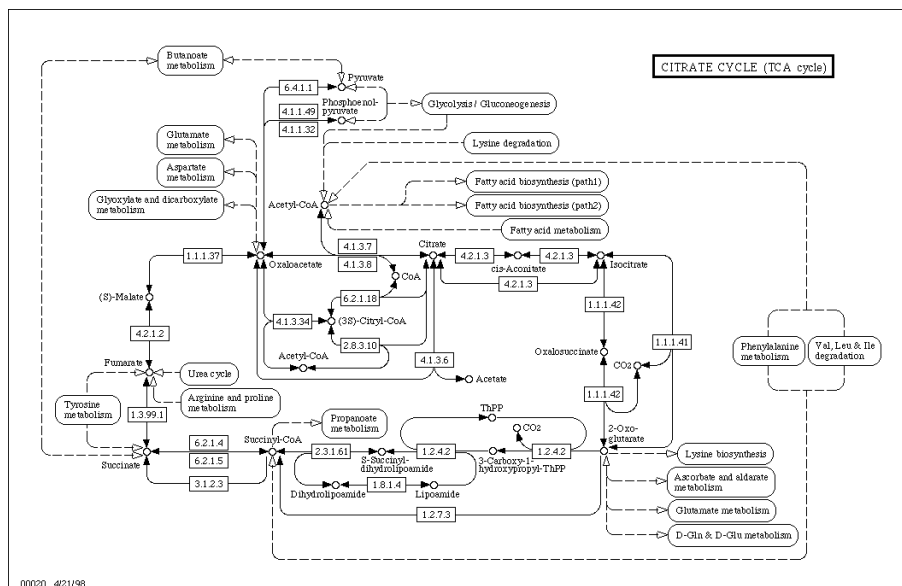


図 4.1: クエン酸回路の PATHWAY マップ
([3] より)

データベース内の 1 エントリーの書式を、以下の図 4.2 に記し、本研究に關与する、4 つのキーの説明を行う。

最上段の ENTRY キーが、前述の EC 番号を表し、酵素の種類を特定している。当データベースは、酵素反応によって分類していると前述したが、分類の対象となる酵素反応に關してのキーが REACTION であり、

REACTION Succinyl-CoA + Dihydrolipoamide = CoA + S-Succinyl-dihydrolipoamide という様に反応式が記述されている。

これらの酵素反応は、代謝反応経路の一部として機能している。その反応が何に關する経路に属しているかを表すキーが、PATHWAY キーである。

PATHWAY PATH: MAP00020 Citrate cycle (TCA cycle)

 PATH: MAP00310 Lysine degradation

という記述は、この酵素反応は、クエン酸回路と、アミノ酸であるリシンを分解する反応系の両方で利用されているという事を意味している。

ここまで述べて来た中で注意すべき事は、ENZYME が扱っている対象は酵素であり、生物種固有の遺伝子では無いことである。酵素と、それぞれの生物種固有の遺伝子との關連は、GENES キーによって、表現している。KEGG 内には、生物種別に既知遺伝子を分

類した GENES データベースがあり、ENZYME データベースの GENES キーは、GENES データベースの ENTRY へリンクされている。

GENES ECO: b0727(sucB)

という記述は、大腸菌 (*Escherichia coli*:ECO) の b0727 遺伝子は、EC 2.3.1.61 という酵素としての機能を持っているという事を示している。また、図 4.2内の GENES キーが指し示す事として、大腸菌 (ECO)、インフルエンザ菌 (HIN)、リケッチア (RPR)、枯草菌 (BSU)、結核菌 (MTU)、トラコーマ病原体 (CTR)、クラミジア (CPN)、ディノコッカス (DRA)、酵母菌 (SCE)、線虫 (CEL) の 10 種類の生物において、EC 2.3.1.61 の酵素が存在しているという事を示す。

```

ENTRY      EC 2.3.1.61
NAME       Dihydrolipoamide S-succinyltransferase
CLASS      Transferases
           Acyltransferases
           Acyltransferases
SYSNAME    Succinyl-CoA: dihydrolipoamide S-succinyltransferase
REACTION    Succinyl-CoA + Dihydrolipoamide = CoA + S-Succinyl dihydrolipoamide
SUBSTRATE   Succinyl-CoA
           Dihydrolipoamide
PRODUCT     CoA
           S-Succinyl dihydrolipoamide
COMMENT     A lipoyl-protein; component of the multienzyme 2-oxoglutarate
           dehydrogenase complex.
PATHWAY     PATH: MAP00020 Citrate cycle (TCA cycle)
           PATH: MAP00310 Lysine degradation
GENES       ECO: b0727(sucB)
           HIN: HI1661(sucB)
           RPR: RP179(sucB)
           BSU: odhB(odhB)
           MTU: Rv2215(sucB)
           CTR: CT055
           CPN: CPn0377
           DRA: DR0083
           SCE: YDR148C(KGD2)
           CEL: W02F12.5
DISEASE     MIM: 126063 Dihydrolipoamide S-succinyltransferase (E2 component
           of
MOTIF       PS: PS00189 [GN]-x(2)-[LIVF]-x(5)-[LIVFC]-x(2)-[LIVFA]-x(3)-K-
           [STAIV]-[STAVQDN]-x(2)-[LIVMFS]-x(5)-[GCN]-x-[LIVMFY]
STRUCTURES  PDB: 1GHJ 1GHK 1HBR 1BAL 1BBL 1E20 1PMR
DBLINKS     University of Geneva ENZYME DATA BANK: 2.3.1.61
           WIT (What Is There) Metabolic Reconstruction: 2.3.1.61
           BRENDA, the Enzyme Database: 2.3.1.61
           SCOP (Structural Classification of Proteins): 2.3.1.61
///

```

図 4.2: ENZYME データベースのエントリー例

EC 4.1.1.32	MAP00020	Citrate cycle (TCA cycle)
	MAP00620	Pyruvate metabolism
EC 2.3.1.61	MAP00020	Citrate cycle (TCA cycle)
	MAP00310	Lysine degradation
EC 4.1.3.6	MAP00020	Citrate cycle (TCA cycle)
	MAP00720	Reductive carboxylate cycle (CO ₂ fixation)
EC 4.1.3.7	MAP00020	Citrate cycle (TCA cycle)
	MAP00630	Glyoxylate and dicarboxylate metabolism
EC 4.1.3.8	MAP00020	Citrate cycle (TCA cycle)
EC 1.8.1.4	MAP00010	Glycolysis / Gluconeogenesis
	MAP00020	Citrate cycle (TCA cycle)
	MAP00260	"Glycine, serine and threonine metabolism"
	MAP00620	Pyruvate metabolism

表 4.2: pathway フィールド切出の例

4.2 クエン酸回路の形成に関する仮説

クエン酸回路の PATHWAY マップ (図 4.1) 内に存在する全酵素の ENZYME データベースから PATHWAY フィールドを抽出した。詳細を付表に示し、例を以下の表 4.2 に示す。例にある表のように、1つの酵素が複数の代謝経路を持つような場合、2つの可能性が考えられる。

- EC 4.1.1.32 は、クエン酸回路 (Citrate cycle) と、ピルビン酸 (Pyruvate metabolism) 系代謝経路の2つの系で働いているように見えるが、KEGGにおける分類上、隣接した経路同士の境界部分に存在する酵素が複数の経路に含まれてしまう。
- EC 4.1.3.6 は、クエン酸回路 (Citrate cycle) と、光合成回路 (Reductive carboxylate cycle) の機能が正反対の系で、働いている。
この場合、どちらか一方の系で機能していた酵素が、他の系で利用されるようになったと考えることが出来る。

クエン酸回路は、生物に必須な代謝系回路ではなく、嫌気性細菌はクエン酸回路を持たない。しかし、回路としては機能していない嫌気性細菌も、部分的にクエン酸回路で使われている酵素を持つことが多い。

本研究では、クエン酸回路のそれぞれの遺伝子は、異なる複数の代謝経路で使われていたものが組み合わせられることによって、クエン酸回路という新しい経路が形成されたという仮説を提起する。

4.3 クエン酸回路遺伝子の分類

クエン酸回路全遺伝子を、ENZYME データベースの PATHWAY フィールドで分類した結果、以下の4つにまとめられた。

- 光合成回路由来の遺伝子
- ピルビン酸、グリオキシル酸、ブタン酸などの炭水化物分解反応系及びエネルギー生成に関する代謝系由来の遺伝子
- アミノ酸の合成及び、分解に関する代謝系由来の遺伝子
- クエン酸回路にしか利用されていない遺伝子

光合成回路	炭水化物	アミノ酸	クエン酸回路のみ
EC 1.1. 1.37	EC 6.4.1. 1	EC 2.3.1.61	EC 4.1.3. 6
EC 4.1. 3. 6	EC 4.1.1.49	EC 1.2.4. 2	EC 4.1.3.34
EC 4.2. 1. 3	EC 4.1.1.32	EC 1.8.1. 4	EC 2.8.3.10
EC 1.1. 1.42	EC 4.1.3. 7		EC 6.2.1.18
EC 1.2. 7. 3	EC 6.2.1. 4		EC 1.1.1.41
EC 6.2. 1. 5			EC 3.1.2. 3
EC 1.3.99. 1			
EC 4.2. 1. 2			

表 4.3: 分類結果

第 5 章

分子進化系統樹による分類結果の検証

本章では、4.3節で分類された酵素群をどの生物が持っているかを示して考察すると共に、分子進化系統樹を作成した。

対象とする生物群は、全塩基配列が決定している 25 種の生物 (図 1.1参照) とした。

マルチプルアライメントを行う際には、BLOSUM アミノ酸置換行列を使用し、系統樹作成アルゴリズムとしては、近接結合法を利用した。また、ブートストラップ確率を求める際には、Kimura の補正式を利用したなお、系統樹は根を仮定していない無根系統樹であるが、表示の便宜上階層的に表現した。

5.1 アミノ酸由来

アミノ酸生合成系由来と思われる遺伝子が、どの生物種にいくつあるかを以下、図 5.1 に示した。示す。アミノ酸生成酵素由来の酵素群は、全て、ケトグルタル酸をスクシニ

	EC 2.3.1.61	EC 1.8.1.4	EC 1.2.4.2
ECO	1	1	1
HIN	1	1	1
HPY	0	0	0
HPJ	0	0	0
RPR	1	2	1
BSU	1	2	1
MGE	0	1	0
MPN	0	1	0
MTU	1	3	1
CTR	1	1	1
CPN	1	1	1
BBU	0	0	0
TPA	0	0	0
SYN	0	1	0
AAE	0	1	0
TMA	0	1	0
MJA	0	1	0
MTH	0	1	0
AFU	0	0	0
PHO	0	0	0
PAB	0	0	0
APE	0	1	0
SCE	1	1	1
CEL	1	1	2
遺伝子数	9	17	9

表 5.1: アミノ酸生成酵素由来と思われる遺伝子の数

ル CoA に酸化するための反応系である。この反応系は、光合成由来の EC 1.2.7.3 の反応に代替する事ができる。次に、これらの酵素群で進化系統樹を作成した。図 5.1 から図 5.2 で示す。

これら 3 つの系統樹には以下の特徴が見受けられる。

3 つの系統樹の中からブートストラップ値の高い個所を抽出すると、(eco,hin)(sce,cel)(ctr,cpn) という 3 つのクラスターが見いだされた。この 3 クラスターを順に A,B,C と略記し、互いの関係を系統樹から記述すると、EC 2.3.1.61 と EC 1.8,1.4 では ((A,B),C)、EC 1.2.4.2 では (A,(B,C)) と異なる結果をだし、そのブートストラップ値は低かった。これらの事から、アミノ酸生合成酵素由来の遺伝子群は、近縁では樹形が保存されている事が確認できた。

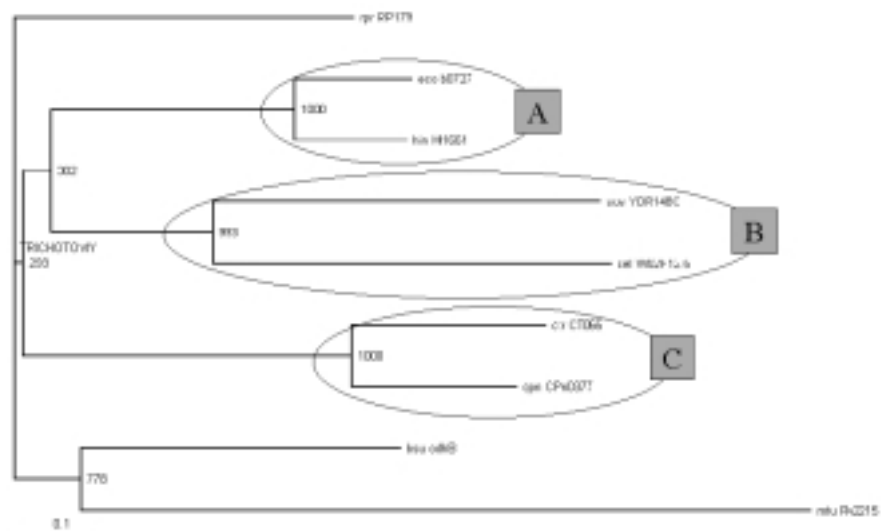


図 5.1: EC 2.3.1.61 遺伝子 (Dihydrolipoamide S-succinyltransferase) の系統樹

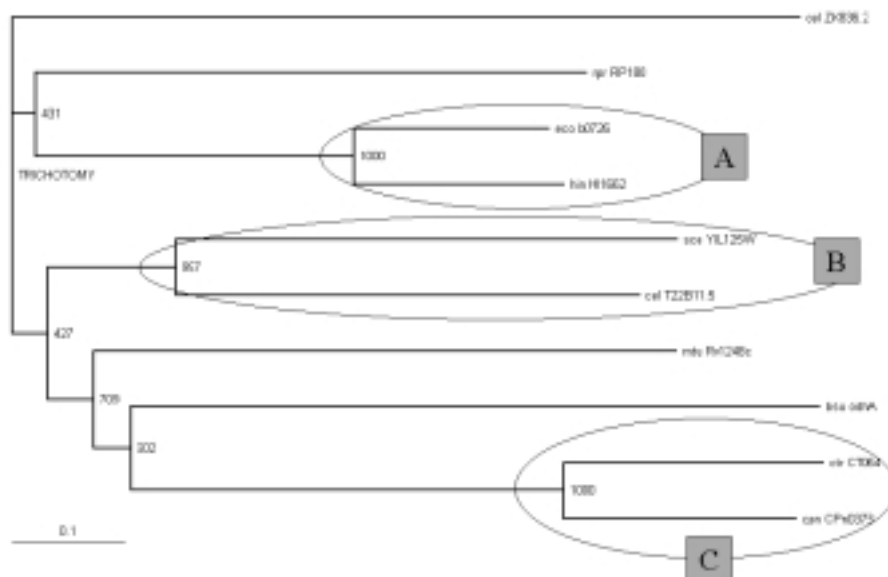


図 5.2: EC 1.2.4.2 遺伝子 (Oxoglutarate dehydrogenase (lipoamide)) の系統樹

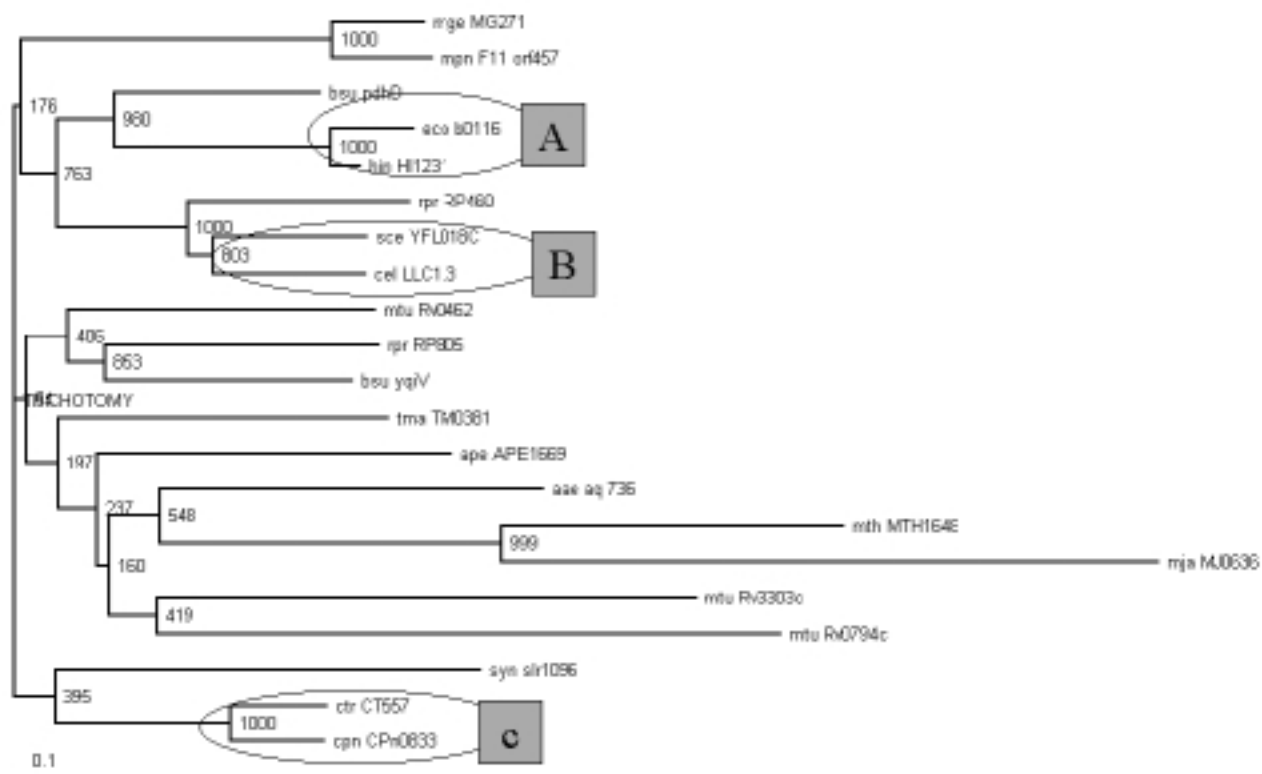


図 5.3: EC 1.8.1.4 遺伝子 (Dihydrolipoamide dehydrogenase) の系統樹

5.2 光合成由来に関する系統樹

光合成回路由来と思われる遺伝子が、どの生物種にいくつあるかを図 5.2に示す。

	EC 1.1.1.37	EC 4.1.3.6	EC 4.2.1.3	EC 1.1.1.42	EC 1.2.7.3	EC 6.2.1.5	EC 1.3.99.1	EC 4.2.1.2
ECO	1	2	2	1	0	2	5	3
HIN	1	2	0	0	0	2	0	1
HPY	0	0	1	1	4	0	0	1
HPJ	0	0	1	1	0	0	0	1
RPR	1	0	1	1	0	2	0	1
BSU	2	0	1	1	0	2	0	1
MGE	0	0	0	0	0	0	0	0
MPN	0	0	0	0	0	0	0	0
MTU	1	1	1	1	0	2	0	1
CTR	1	0	0	0	0	2	0	1
CPN	1	0	0	0	0	2	0	1
BBU	0	0	0	0	0	0	0	0
TPA	0	0	0	0	0	0	0	0
SYN	1	0	1	1	0	2	0	1
AAE	2	0	1	1	0	4	0	2
TMA	0	0	0	1	0	0	0	2
MJA	2	0	0	2	0	2	0	2
MTH	2	0	0	1	4	2	0	4
AFU	1	0	0	1	4	4	0	2
PHO	1	0	0	0	0	0	0	2
PAB	1	0	0	0	0	0	0	2
APE	1	1	1	1	0	2	0	1
SCE	3	0	2	3	0	0	0	1
CEL	1	1	2	2	0	1	0	1
遺伝子数	17	5	11	15	3	14	1	20

表 5.2: 光合成回路由来で系統樹作成に利用した数

光合成回路由来と思われる遺伝子群は、クエン酸回路の根幹をなしていて、数量的にも多いが、全ての遺伝子群を満たしているような生物種はない。クエン酸回路が機能していない生物群では当然の事であるが、クエン酸回路を所持している真核生物においても、いくつかの欠が見える。

酵母 (SCE)、線虫 (CEL) で共に欠している EC 1.3.99.1 は、大腸菌以外では登録されていない。考えられる可能性としては、機能未知遺伝子の中に該当する遺伝子が含まれることや、他の酵素で反応が代替されている事などがあげられる。

酵母で欠けている EC 4.1.3.6 が誘発する反応は、炭水化物分解系由来と分類した EC 4.1.3.7 で代替されている。

EC 1.2.7.3 と EC 6.2.1.5 というひと続きの遺伝子は、アミノ酸由来と分類した EC 2.3.1.61, EC 1.8.1.4, EC 1.2.4.2 と、炭水化物分解系由来に分類した EC 6.2.1.4 によって反応が代替されている。

線虫で欠けている 1.2.7.3 の反応も、酵母と同様アミノ酸由来三酵素によって代替されている。

次に、これらの酵素群で進化系統樹を作成した。図 5.5から 5.12に示した。アミノ酸生

成系の様な全酵素で共通な、生物種クラスターは確認できなかった。また、系統樹内でのブートストラップ値が低い傾向が見いだされた。(図 5.4参照)

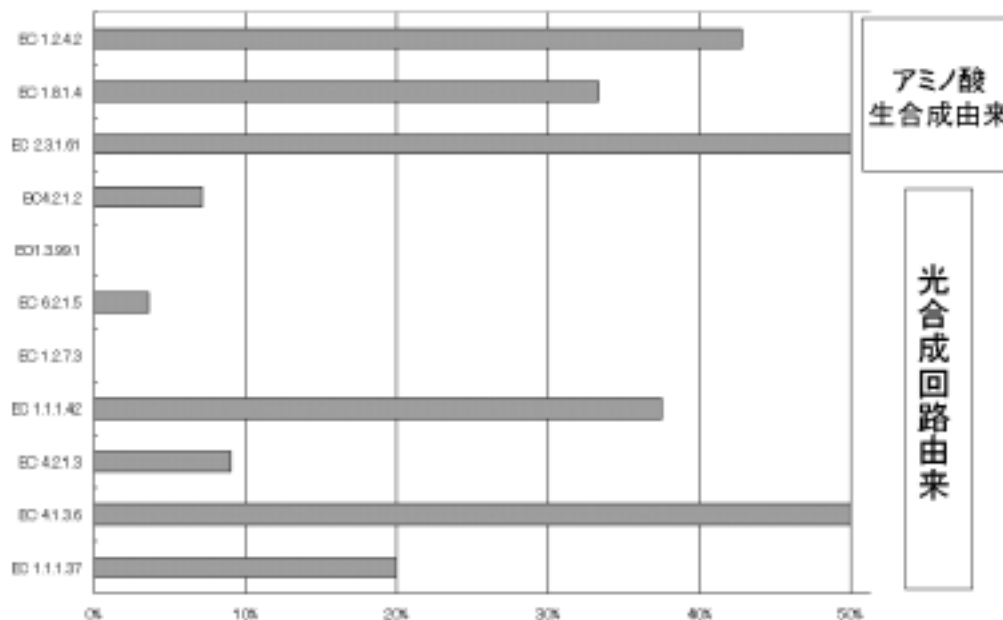


図 5.4: 系統樹内の全ブートストラップ値の中で、95%以上である確率

このグラフは、系統樹内の全ブートストラップ値が、95%以上である確率を表している。アミノ酸由来遺伝子はブートストラップ値が高いクラスターの割合が多いのに対し、光合成回路由来遺伝子は、割合が低いものが目立つ。さらに、光合成回路由来の中で高い割合を示している、EC 4.1.3.6 と EC 1.1.1.42 には注意が必要である。EC 4.1.3.6 は、標本数が3と少数な中に、同じ生物種の同系統遺伝子が重複しているため、ブートストラップ値が高くなっていると考えられる。EC 1.1.1.42 は、標本数も多く信頼性が高い。この酵素の他の機能として、アミノ酸系化合物であるグルタチオン生成を担っていることから、アミノ酸由来と光合成由来の比較の際には妥当でない。

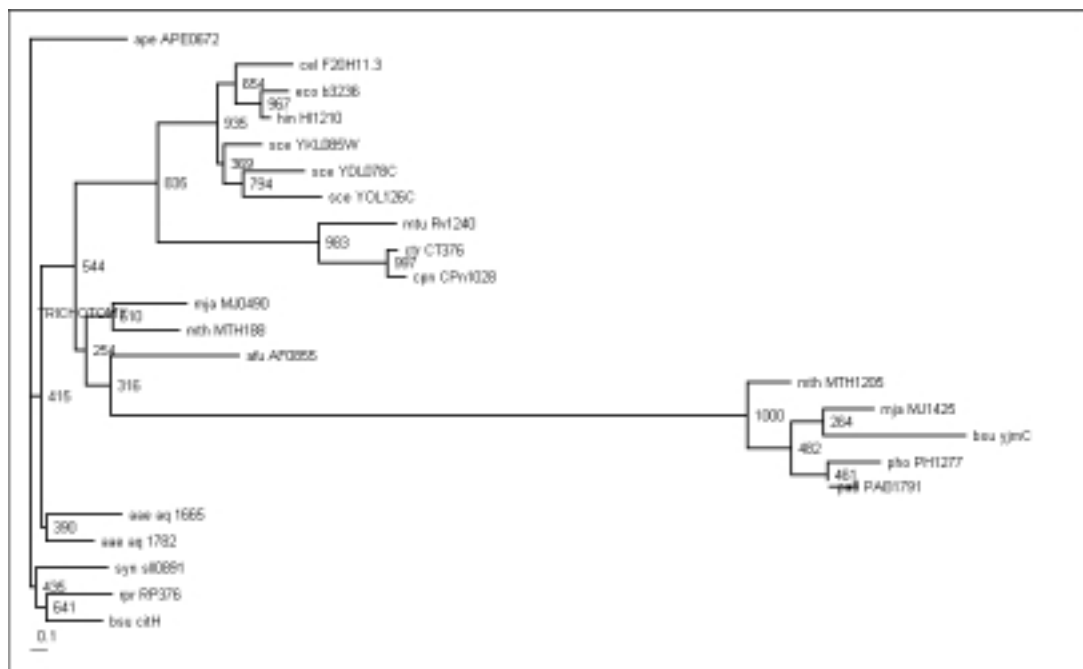


図 5.5: EC 1.1.1.37 遺伝子 (Malate dehydrogenase) の系統樹

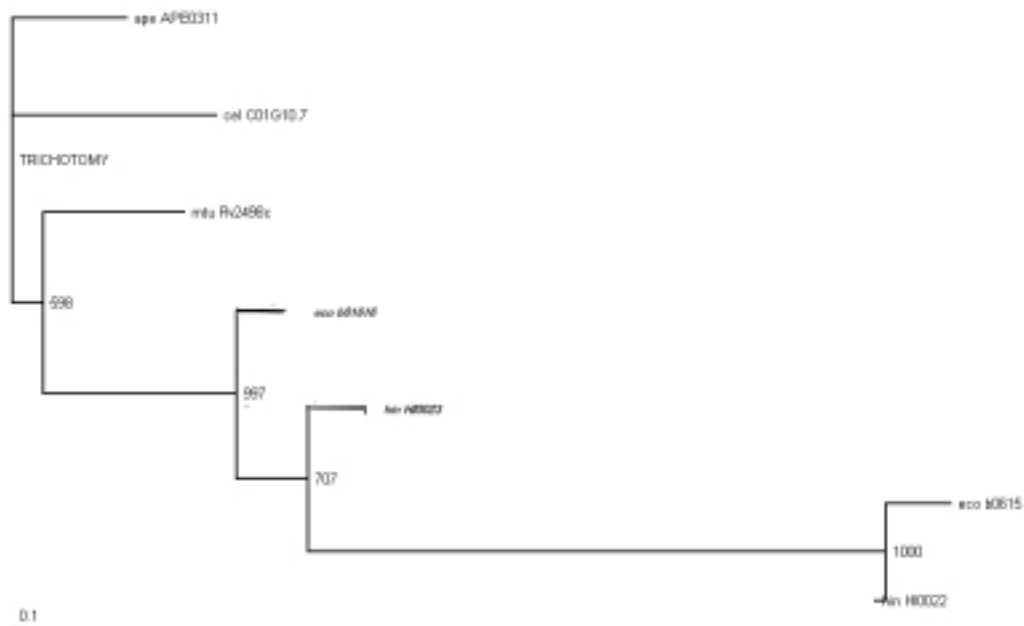


図 5.6: EC 4.1.3.6 遺伝子 (Citrate lyase) の系統樹

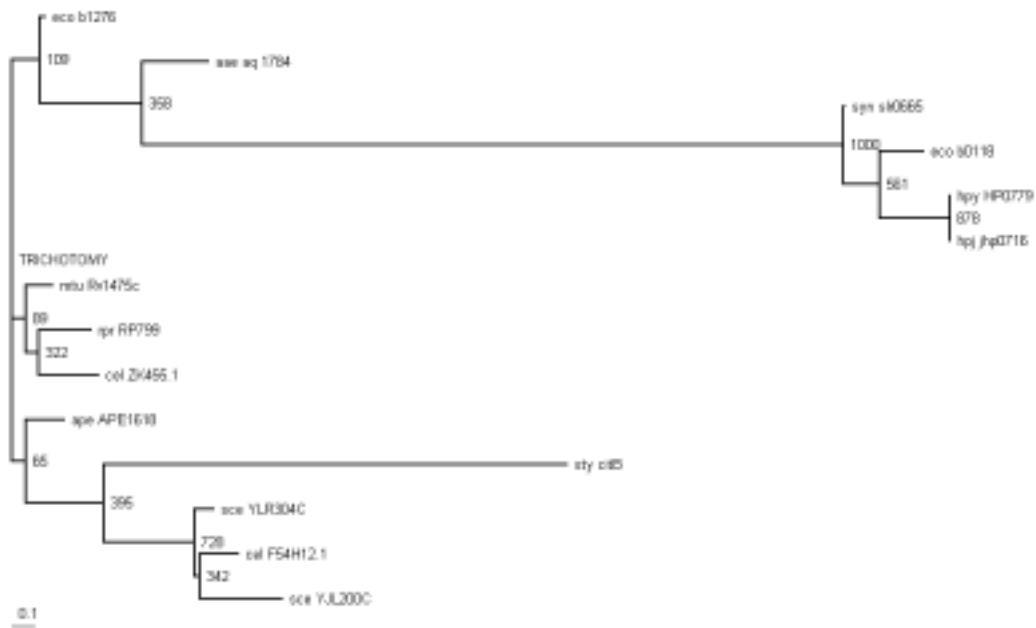


図 5.7: EC 4.2.1.3 遺伝子 (Aconitate hydratase) の系統樹

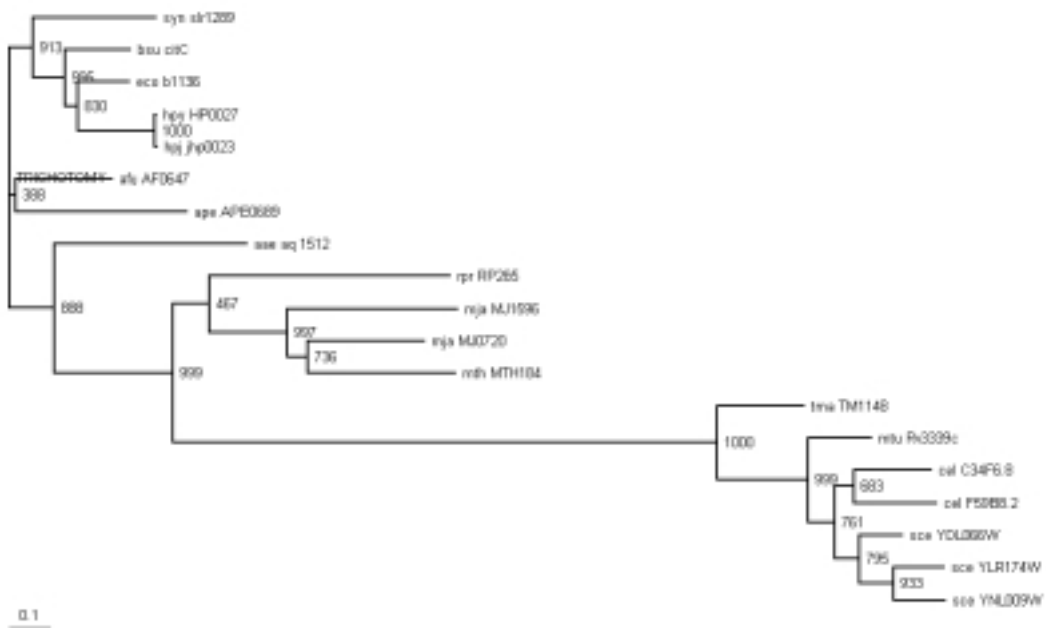


図 5.8: EC 1.1.1.42 遺伝子 (Isocitrate dehydrogenase (NADP+)) の系統樹

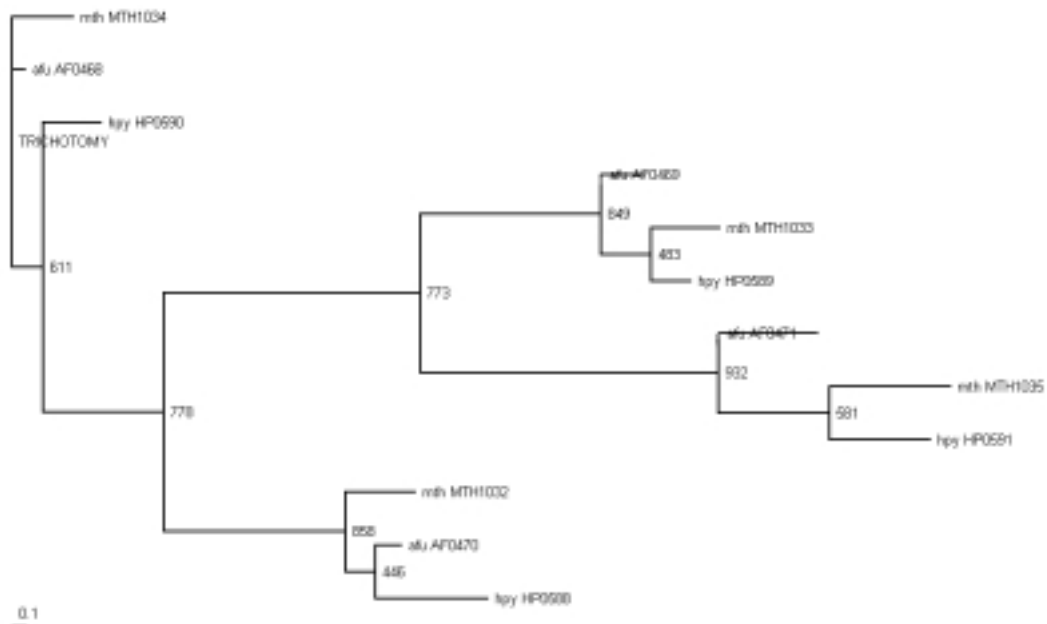


図 5.9: EC 1.2.7.3 遺伝子 (2-Oxoglutarate synthase) の系統樹

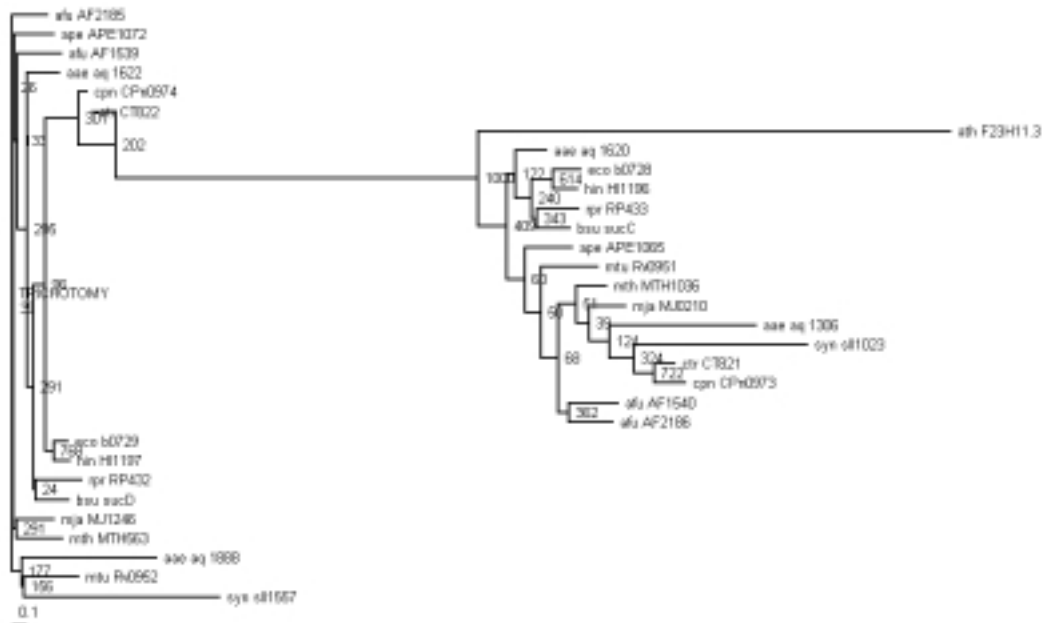
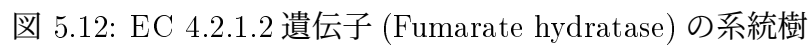


図 5.10: EC 6.2.1.5 遺伝子 (Succinate-CoA ligase (ADP-forming)) の系統樹



5.3 炭水化物分解系由来

炭水化物分解系由来と思われる遺伝子が、どの生物種にいくつあるかを以下に示す。

	EC 6.4.1.1	EC 4.1.1.49	EC 4.1.1.32	EC 4.1.3.7	EC 6.2.1.4
ECO	0	1	0	1	0
HIN	0	1	0	0	0
HPY	0	0	0	1	0
HPJ	0	0	0	1	0
RPR	0	0	0	1	0
BSU	1	1	0	3	0
MGE	0	0	0	0	0
MPN	0	0	0	0	0
MTU	1	0	1	3	0
CTR	0	0	1	0	0
CPN	0	0	1	0	0
BBU	0	0	0	0	0
TPA	0	0	0	0	0
SYN	0	0	0	1	0
AAE	1	0	0	1	0
TMA	0	0	0	1	0
MJA	0	0	0	0	0
MTH	0	0	0	2	0
AFU	0	0	0	1	0
PHO	0	0	1	0	0
PAB	0	0	1	0	0
APE	0	0	0	1	0
SCE	2	1	0	3	2
CEL	1	0	2	0	3
遺伝子数	6	4	7	20	5

表 5.3: 炭水化物分解系由来と推測された遺伝子の数

炭水化物分解系由来の酵素群は、さらに2つに分類することが出来る。

一つ目は、EC 6.4.1.1、EC 4.1.1.49、EC 4.1.1.32 が属す、ピルビン酸もしくはリン酸化ピルビン酸を、オキサロ酢酸にする反応群で、これらは、光合成由来の酵素群と機能が重複しない反応である。

二つ目は、EC 4.1.3.7、EC 6.2.1.4 で、それぞれ、光合成回路由来の EC 4.1.3.6、EC 6.2.1.5 の機能と重複しており、代替可能である。次に、これらの酵素群で進化系統樹を作成した。図 5.13から図 5.17に示した。この酵素群を持っている種は片寄りがあるため、どの酵素にも存在する種のクラスターは発見できなかった。

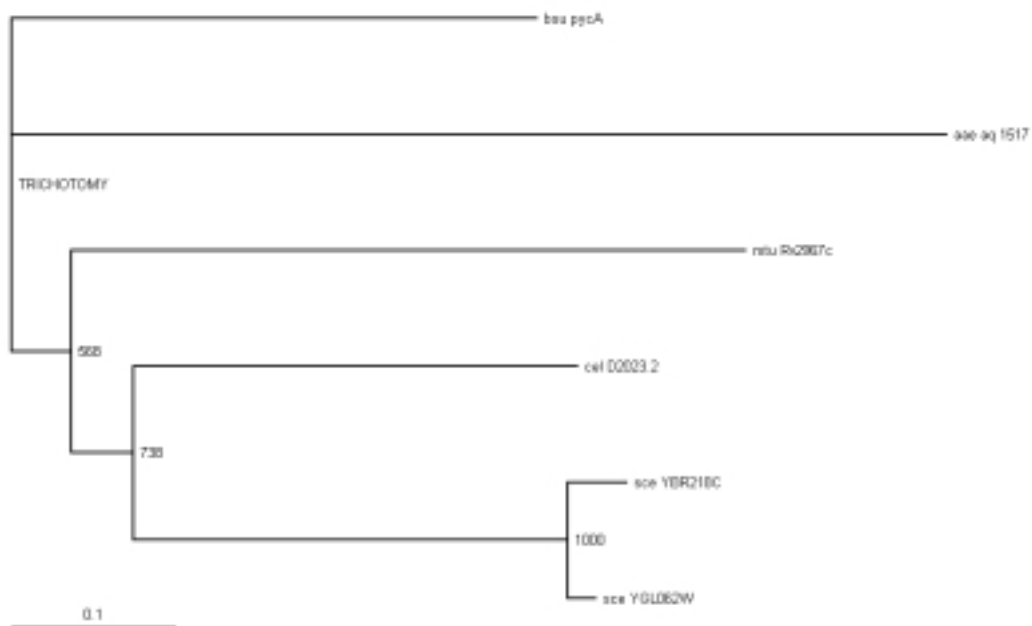


図 5.13: EC 6.4.1.1 遺伝子 (Pyruvate carboxylase) の系統樹



図 5.14: EC 4.1.1.49 遺伝子 (Phosphoenolpyruvate carboxykinase (ATP)) の系統樹

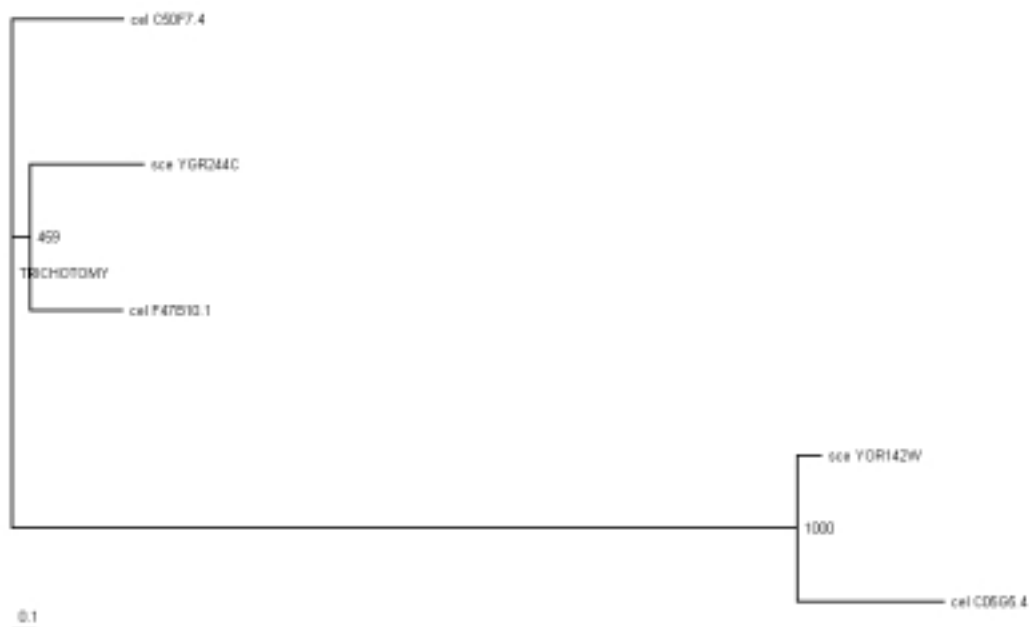


図 5.17: EC 6.2.1.4 遺伝子 (Succinate-CoA ligase (GDP-forming)) の系統樹

5.4 クエン酸回路固有な酵素群

以下が生物種が、クエン酸回路にしかない遺伝子をいくつ所持しているかを表した表である。クエン酸回路独自の遺伝子を持つ生物種は、EC 4.1.3.8 では、線虫 1 種類、EC 1.1.1.41 では線虫と酵母の 2 種類、EC 4.1.3.34 と EC 2.8.3.10 では、インフルエンザ菌の 1 種類のみであった。

EC 4.1.3.8、EC 4.1.3.34、EC 2.8.3.10 は、オキサロ酢酸とアセチル CoA をクエン酸に変化するとという行程を、光合成回路由来と考えられる EC 4.1.3.6 とは別の行程で実行しているため、種に固有な進化を遂げたものと考えられる。また、EC 1.1.1.41 は、イソクエン酸をケトグルタル酸に変換する反応であるが、この酵素も種で独自に進化してきたものと思われる。なお、どの酵素も、3 種未満の生物にしか見つからなかったため、系統樹の作成は行わなかった。

	EC 4.1.3.8	EC 1.1.1.41	EC 4.1.3.34	EC 2.8.3.10
ECO	0	0	0	0
HIN	0	0	1	1
HPY	0	0	0	0
HPJ	0	0	0	0
RPR	0	0	0	0
BSU	0	0	0	0
MGE	0	0	0	0
MPN	0	0	0	0
MTU	0	0	0	0
CTR	0	0	0	0
CPN	0	0	0	0
BBU	0	0	0	0
TPA	0	0	0	0
SYN	0	0	0	0
AAE	0	0	0	0
TMA	0	0	0	0
MJA	0	0	0	0
MTH	0	0	0	0
AFU	0	0	0	0
PHO	0	0	0	0
PAB	0	0	0	0
APE	0	0	0	0
SCE	0	2	0	0
CEL	2	3	0	0
遺伝子数	2	5	1	1

表 5.4: クエン酸回路固有と推定された遺伝子数

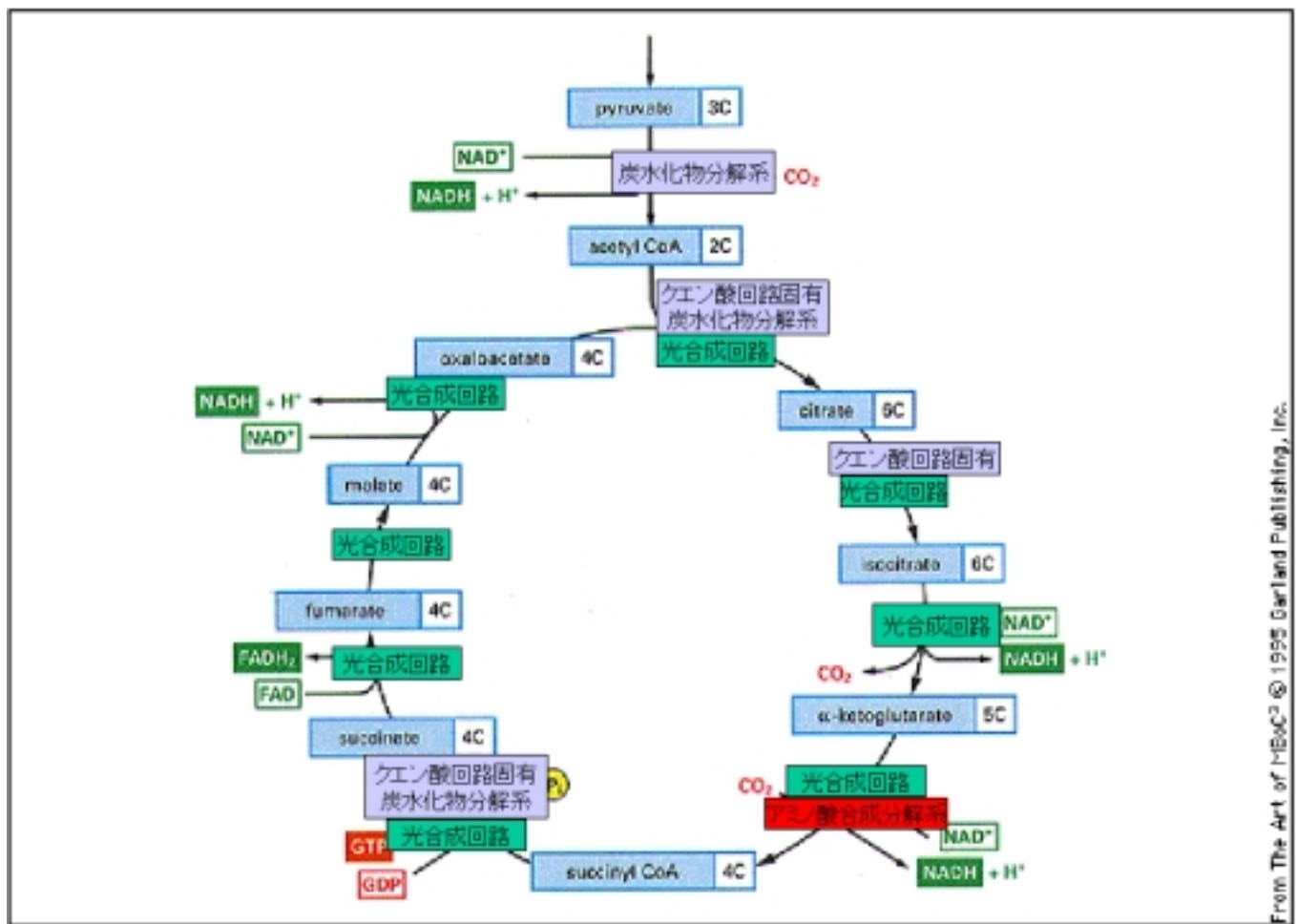
5.5 まとめ

第 4 章と第 5 章では、データベースからクエン酸の由来について仮説を立てる部分と、その結果から系統樹を作成する 2 段階を行って来た。

データベースから由来を求める部分では、クエン酸回路を、光合成回路由来、アミノ酸分解合成系、炭水化物分解系、分解系、クエン酸回路に固有な遺伝子の 4 段階に分けることができた。光合成回路由来に関しては、クエン酸回路は反応系が同じであることから容易に推測できる (図 5.18 参照)。アミノ酸由来遺伝子は、ケトグルタル酸をスクシニル CoA にする反応の周辺に固まっていた。炭水化物分解系由来と思われる遺伝子群は解糖系に近いところに多くみうけられた。クエン酸回路固有の遺伝子の多くは、オキサロ酢酸とアセチル CoA をクエン酸にする周辺の反応に多く見受けられた。この部分は、クエン酸回路を循環させる特異的な部分であることが原因ではないかと思われる。

系統樹作成からは、以下の特徴が認識された。アミノ酸合成系の由来の遺伝子は、近縁では樹形を保存していて、信頼性も高かった。これに対し、光合成回路由来の遺伝子は、樹形が保存されておらず、ブートストラップ確率も低く信頼性に欠けていた。この理由としては、以下のように解釈した。

今回対象とした生物種は全ゲノム配列を持っている種で、クエン酸回路の機能を保持しているものは少数である。その中で、クエン酸回路の反応の中枢を占めている光合成回



From The Art of MBoC © 1995 Garland Publishing, Inc.

図 5.18: 4 分類のクエン酸回路の中での位置

路由来の遺伝子は共通に存在する必要がない。これに対し、アミノ酸合成系由来の遺伝子は、その本来持っていたアミノ酸合成という機能により、樹の形が保存していたものと思われる。

考察

		生物種																							
		大腸菌			インフルエンザ			枯草菌			クラミジア			酵母			精虫								
		ECO	HBI	HPY	HPJ	MCE	MPN	NTU	CTR	CPN	BBU	TPA	SYN	ARE	TMA	MJA	MTH	AFU	PHD	PAB	APL	SCC	CEL		
光合成細菌	光合成細菌	ECO13.58	1	1	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	1	1	
	光合成細菌	ECO13.4	1	1	0	0	1	1	0	1	1	0	0	1	1	1	1	1	0	0	0	1	1		
	光合成細菌	ECO12.42	1	1	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	1	1	
	光合成細菌	ECO13.23	1	1	0	0	0	0	1	1	1	0	0	1	2	0	2	2	1	1	1	1	3	1	
	光合成細菌	ECO13.24	2	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	4	1	
	光合成細菌	ECO13.3	2	0	1	1	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	1	2	1	
	光合成細菌	ECO13.42	1	0	1	1	0	0	1	0	0	0	0	1	1	1	2	1	1	0	0	1	3	1	
	光合成細菌	ECO12.73	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	4	4	0	0	0	0	0	1
	光合成細菌	ECO13.3	2	2	0	0	0	0	2	2	2	0	0	2	4	0	2	2	4	0	0	2	1	1	1
	光合成細菌	ECO12.981	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
光合成細菌	ECO13.3	3	1	1	1	0	0	1	1	1	0	0	1	2	2	2	4	2	2	2	1	1	1	1	

図 6.1 から、アミノ酸由来遺伝子群 3 種全てをもっている生物種は、eco(大腸菌)、hin(インフルエンザ菌)、MTU(結核菌)、CTR(クラミジア)、CPN(クラミジア) の 5 種類である。さらに、この 5 種類の生物種が持っている光合成由来遺伝子群は、EC1.1.1.37、EC6.2.1.5、EC4.2.1.2 であるため、光合成由来遺伝子の系統樹比較にはこの 3 種の遺伝子を採用した。

これらの生物種の遺伝子群の系統樹比較を行う前に、Triose-phosphate isomerase(以下 TPI 遺伝子と略) の系統樹を作成した。(図 6.2参照)

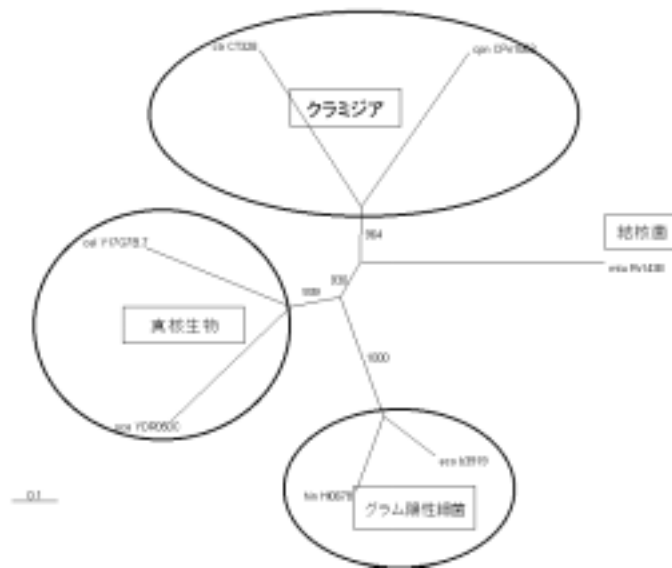


図 6.2: Triose-phosphate isomerase 遺伝子の系統樹

TPI 遺伝子は解糖系で重要な役割を果たす遺伝子であり、その由来は古い。その為、遺伝子系統樹は信頼性が高い事が知られている。この事を利用して、本研究では TPI 遺伝子を模範データとして使用した。系統樹の特徴としては、 $((eco,hin)(sce,cel))mtu,(ctr,cpn))$ となった。この結果は、グラム陽性細菌、真核生物、クラミジアという近縁生物種が同一クラスターを形成し、それぞれのブートストラップ値も高く信頼性が高かった。本章では、以降 TPI の系統樹を分子進化のモデル系統樹として用いる。

6.1 光合成由来遺伝子の系統樹比較

以下に、光合成由来遺伝子 3 種の系統樹を示す。

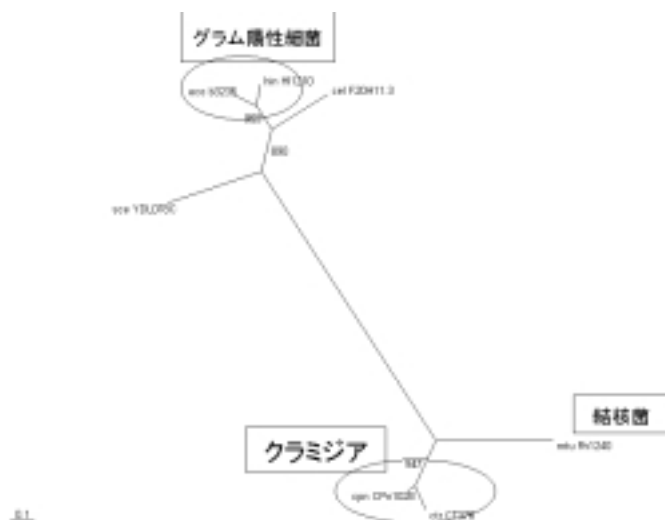


図 6.3: EC1.1.1.37 遺伝子 (Malate dehydrogenase) の系統樹

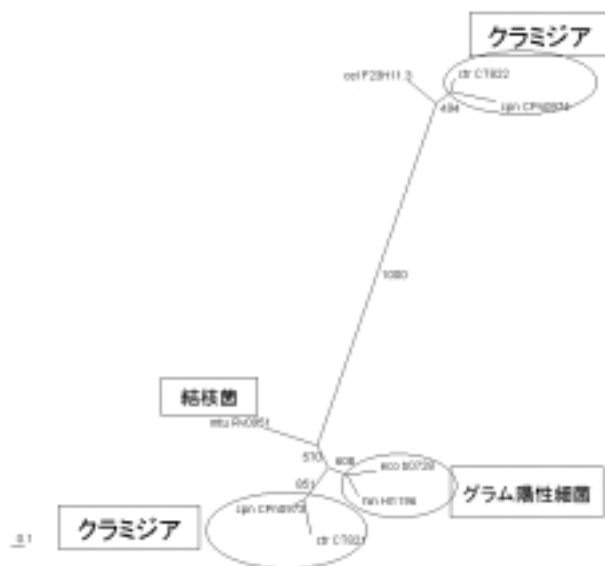


図 6.4: EC6.2.1.5 遺伝子 (Succinate-CoA ligase (ADP-forming)) の系統樹

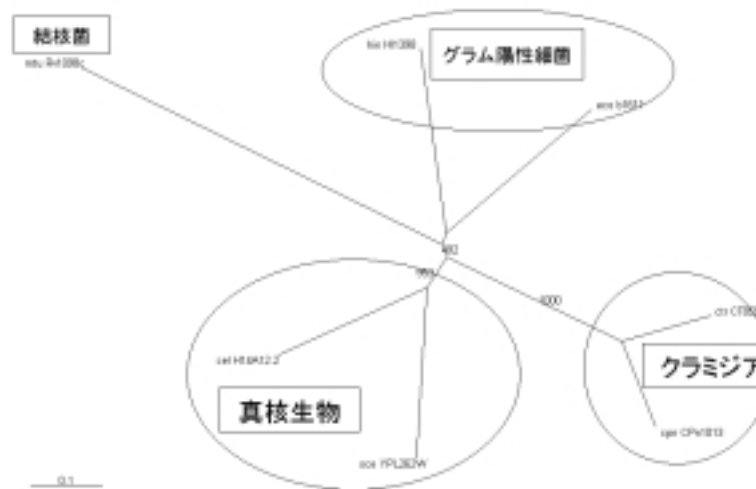


図 6.5: EC4.2.1.2 遺伝子 (Fumarate hydratase) の系統樹

これらの系統樹は、グラム陽性細菌、クラミジア、真核生物といった近縁のクラスターは高頻度で形成しているが、クラスター間のトポロジーはそれぞれ異なる。また、後述するアミノ酸由来遺伝子群の系統樹に比べてブートストラップ値が低いことも分かった。

6.2 アミノ酸由来遺伝子の系統樹比較

以下に、アミノ酸由来遺伝子 3 種の系統樹を示す。

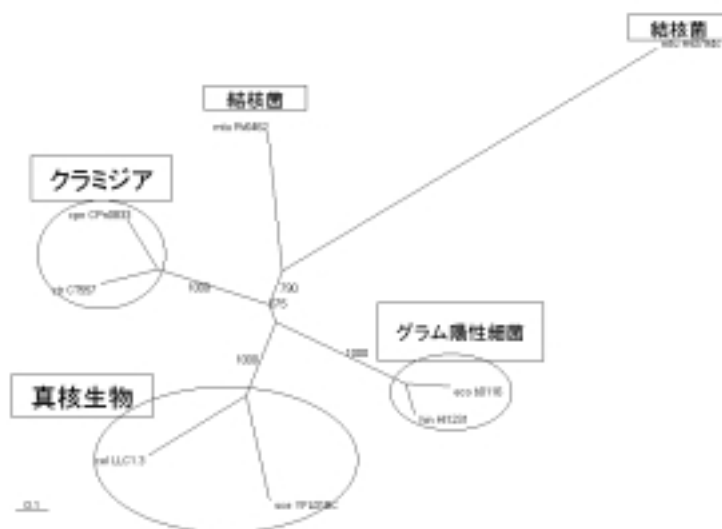


図 6.6: EC1.8.1.14 遺伝子 (Dihydrolipoamide dehydrogenase) の系統樹

アミノ酸由来遺伝子の系統樹は3種とも同一なトポロジーを示し、TPI 遺伝子の系統樹のトポロジーとも等しかった。また、光合成由来遺伝子の系統樹に対して、高いブートストラップ値を示した。このことから、アミノ酸生合成系由来の3遺伝子の由来は古く、系統樹の作成に適している事が分かった。

6.3 アライメント比較

本章では、アミノ酸由来遺伝子は全ての遺伝子においてトポロジーが等しく系統樹の信頼性が高いが、光合成由来遺伝子は、トポロジーが保存されておらず、系統樹の信頼性が低いことが分かった。本節では、この理由について考察を行う。以下、アミノ酸由来遺伝子である EC2.3.1.61 の遺伝子系統樹の作成で使われたアライメントを図 6.3、光合成回路由来遺伝子である EC4.2.1.3 の遺伝子系統樹の作成で使われたアライメントを図 6.3に示す。

```
eco_b0727      LRKRAERLLEAKNSTAMLTTFNEVNMKPIMDLRKQYGEAFEKRHGIRLGFMSFYVKAVV
hin_HI1661     LRKRIAERLLEAKNSTAMLTTFNEVDMQPIIMTLRKTYGEKFEKQHSVRLGFMSFYIKAVV
sce_YDR148C    MRLRIAERLKESQNTAASLTTFNEVDMSALMEMRKLYKDEIIKKTGTFKFGMGLFSKACT
cel_W02F12.5   MRMRIAQRLKDAQNTYAMLTTFNEIDMSSLIEMRKTYQKDFVAKHGVLGMMSPFVRAAA
ctr_CT055      IRKTISRRLVQSLHDSAMLTTFNEIHMGPLIALRKERQEDFVAKYGVKLGMSFFVRAVV
cpn_CPn0377    IRKTISRRLLSALHESAMLTTFNEVYMTPLFHLRKEKQEELSRYGVKLGMSFFVKAVL
mtu_Rv2215     IRQITANKTRESLQATAQLTQTHEVDMTKIVGLRARAKAAFAEREGVNLTFLPFFAKAVI
               :*  .:.  .:  :  *  **  :*  *  .:  .:*      :  :  .  .:  :  :  :  :*

eco_b0727      EALKRYPEVNASIDGD--DVVYHNYFDVSMVSTPRGLVTPVLRDVTLGMDIEKKIKE
hin_HI1661     EALKRYPEVNASIDGD--DVVYHNYFDISIAVSTPRGLVTPVLRDCDKLSMAEIEKQIKA
sce_YDR148C    LAAKDIPAVNGAIEGD--QIVYRDYDISVAVATPKGLVTPVVRNAESLSVLDIENEIVR
cel_W02F12.5   YALQESPVVNAVLDEN--EIVYRHFVDISVAVATPKGLVVPVLRNVESMNYAQIELELAN
ctr_CT055      DSLKKYPRVNAYIEDN--EIVYRHYYDISIAIGTDRGLVVPVIRNCDQLSSGEIELQLAD
cpn_CPn0377    EALKAYPRVNAYIDGE--EIVYRHYYDISIAVGIDRGLVVPVIRDCDKLSNGEIEQKLAD
mtu_Rv2215     DALKIHPNINASYNEDTKEITYDAEHLGFAVDTEQGLLSPVIRDAGDLSLAGLARAIAD
               :  :  *  :*,  :  :  :  :*,  .  .:.*,  :  :*,  :  :*,  :  :  :

eco_b0727      LAVKGRDGKLTVEDLTGGNFTITNGGVFGSLMSTPIINPPQSAILGMHAIKDRP-----M
hin_HI1661     LAEKGRDGKLTVEDLTGGNFTITNGGVFGSLMSTPIINPPQSAILGMHAIKERP-----I
sce_YDR148C    LSHKARDGKLTLEDMTGGTFTISNGGVFGSLYGTPIINSPQTAVLGLHG VKERP-----V
cel_W02F12.5   LGVKARDGKLAVEDMEGGTFTISNGGVFGSMFGTPIINPPQSAILGMHGP EIR-----V
ctr_CT055      LASRAREGKLAIHELEGGGFTITNGGVYGSLLSTPIINPPQVGILGMHKIEKRP-----V
cpn_CPn0377    LALRAREGLLAI AELEGGGFTITNGGVYGSLLSTPIINPPQVGILGMHKIEKRP-----V
mtu_Rv2215     IAAARARSGNLKPDELSGGTFTITNIGSQALFDTPILVPPQAAMLGTAIVKRPRVVVDA
               :.  :  :*,*  *  :  :  **  :*,*  *  *  :  :  :*,*  :  :*,*  :  :*,*  *

eco_b0727      AVNGQVEILPMMYLALS YDHRLIDGRESVGFLVTIKELLED P
hin_HI1661     ALNGQVVIRPMMYLALS YDHRLIDGRESVGFLVTIKELLED P
sce_YDR148C    TVNGQIVSRPMMYLALTYDHRLLDGREAVTFLKTVKELIED P
cel_W02F12.5   -----PIMQIALTYDHRLIDGREAVTFLKKIKTAVEPD
ctr_CT055      VREDAIVIADMMYVAMS YDHRIIDGKEAVGFLVNVKEQLEQ P
cpn_CPn0377    VLDNEIVIADMMYVALS YDHRLIDGKEAVGFLVKVKEGLENP
mtu_Rv2215     SGNESIGVRSVCYLPLTYDHRLIDGADAGRFLT TTIK HRL EEG
               :  :  :*,**,*  :  :  **  .,*  :  :*
```

図 6.9: EC2.3.1.61 遺伝子 (アミノ酸生合成由来) のアライメント

```

sce_YLR304C      IHQIVLENYAFP GAL----- IIGTDSHTPNAGGLGQLAIGVGGADAVDVMAGR PW
cel_F54H12.1     IHQIILENYAFP GLL----- LIGTDSHTPNNGGLGGLCIGVGGADAVDMADIPW
sce_YJL200C      IHQIVLENFSAPGLM----- MLGTDSTPNAGGLGAIAGVGGADAVDALTGTPW
mtu_Rv1475c      VHQVNI EYLASV VMT-RDGVAYP DTCVGTDSHTTMV NGLGVLGWGVGGIEAE AMLGQP V
cel_ZK455.1      VHQVNI EYLART VFVGK DGVLPDSVVG TDSHTT MIDSGVLGWGVGGIEAE AVM LGQP I
eco_b0118        LHALAMLKNAREGIEPDQPGVVG--PIKQIEALQQKGFPLAYVGDVVG TGSSRKSATNSV
                  : * : : : : : : : : : : : : : : : : : : : : : : : : : : :
                  : * : : : : : : : : : : : : : : : : : : : : : : : : : : :

sce_YLR304C      ELKAPKILGVKLTGKMNGWTS PKDIILKLAGITTVKGGTGKIVEYFGDGVDTFSATGMGT
cel_F54H12.1     ELKCPKVIGIKLTGKLNWTS AKDVILKVADILTVKGGTG AIVEYFGPGVDSISATGMGT
sce_YJL200C      ELKAPKILGVKLTGKLNWSTPKDVITKLAGLLTVRGGTGYIVEYFGEGVSTLSCTGMAT
mtu_Rv1475c      SMLIPRVVGFRLTGEIQPGVTA TDVVLTVTEMLRQHGVVGK FVEFYGEGVAE VPLANRAT
cel_ZK455.1      SMVPIEVIGYELVGTLSDTVTS TDLVLTITKNLRDLGVVGK FVEFFGTGVASLSIADRAT
eco_b0118        LWFMGDDIPHPV NKRGGGLCLGGKIAPIFFNTMEDAGALPIEVDVSNLNMGDVIDVYPYK
                  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
                  : : : : : : : : : : : : : : : : : : : : : : : : : : : :

sce_YLR304C      ICNMGAIEGATTSVFPFNKSMIEYLEATGRG-----KIADFAKLYH--KDLSADKDAEY
cel_F54H12.1     ICNMGAIEGATTSVFPYNESMYKYLEATGRK-----EIAEEARKY---KDLLTADDGANY
sce_YJL200C      ICNMGAIEGATTSTFPYQEAHKRYLQATNRA-----EVAEADVALNKFNFLRADKDAQY
mtu_Rv1475c      LGNMSPEFGSTA AIFPID EETIKYLRFTGRTP EQVALVEAYAKAQGMWH---DPKHEPEF
cel_ZK455.1      IANMCPEYGATIGFFPVDSRTIDYLTQTGRD TDYTRVEQYLKSVGMFVNFTDDSYRPTY
eco_b0118        GEVRNHETGELLATFELKTDVLI DEVRAGGRIP LIIGRLTTKAREALGLPHSDVFRQAK
                  * * . * . : .
                  : .

sce_YLR304C      DEVVEIDLNTLEPYINGPFTPD L ATPVSKMKEVAVANNW
cel_F54H12.1     DQIIEINLDTLTPHVNGPFTPD LASSIDKLG ENAKKNW
sce_YJL200C      DKVIEIDLSAIEPHVNGPFTPD LSTPISQYAEKSLKENW
mtu_Rv1475c      SEYLELNLSDVVPSIAGPKRPQDRIALAQAKS TFREQIY
cel_ZK455.1      TTTLKLDLGSVVPSVSGPKRPHDRVELASLAQDFSKGLT
eco_b0118        DVAESDRGFSLAQKMVGRACGVKGIRPGAYCEPKMTSVG
                  . : : * .

```

図 6.10: EC4.2.1.3 遺伝子 (光合成回路由来) のアライメント

この両アライメントを比較すると、EC2.3.1.61 遺伝子 (アミノ酸生合成由来) のアライメントは一致率が高く配列が保守的であるのに対し、EC4.2.1.3 遺伝子 (光合成回路由来) のアライメントは一致率が低く、分子進化が起こりやすい事が分かった。

この理由として、光合成由来遺伝子は現在は光合成回路の機能を失いクエン酸回路の機能のみしか果たしていない為、機能的制約が緩く突然変異が固定されやすい。これに対し、アミノ酸由来遺伝子は、現在でもアミノ酸生合成機能とクエン酸回路機能を担っており、機能的制約がきつく、突然変異が固定されにくく、配列は保守的だと推論した。

第 7 章

結語

本研究では、酵素データベースから酵素の由来を推定し、その結果を分子進化系統樹の形状の違いを検証した。

この結果、クエン酸回路遺伝子は、光合成回路遺伝子由来、アミノ酸生合成遺伝子由来、炭水化物生合成遺伝子由来、クエン酸回路固有な遺伝子の 4 種に分類され、アミノ酸生合成由来遺伝子は系統樹の信頼性が高く、系統樹の形状も一致した。

以上のことから、遺伝子系統樹の推定を行う際には、遺伝子の由来、現存する機能両面から考えることが有用である事が分かった。また、古い分岐を考える上では複数の代謝経路で機能している遺伝子から推定する事が適当である事が分かった。

謝辞

本研究は、小長谷明彦教授の御指導、御助言のもと行われました。ここに深く感謝の意を表します。

遺伝子知識システム論、佐藤賢二助教授には、多くの御助言を頂き、貴重な時間を割いて頂きました。ここに謝意を表します。

また、研究遂行にあたり、多くの協力を頂いた遺伝子知識システム論講座の講座の皆様、感謝の意を表します。

参考文献

- [1] Edward M.Marcotte,Matteo Pellegrini,Michael J.Thompson,Todd O.Yeates &David Eisenberg,A combined algorithm for genomewide prediction of protein function,nature volume 402 p83-p86—issue no.6757,4 November 1999
- [2] Anton J.Enright,Ioannis Iliopoulos,Nikos C.Kyrpides & Christos A.Ouzounis, Protein interaction maps for complete genomes based on gene fusion events,nature volume 402 p86-p90—issue no.6757,4 November 1999
- [3] Kyoto_Encyclopedia_of_Gene_and_Genomes(KEGG)
<http://www.genome.ad.jp/kegg/kegg2.html>
- [4] R.Durbin,S.Eddy,A.Krogh,G.Mitchison, Biological sequence analysis,Cambridge University Press.1998
- [5] P.Baldi,S.Brunak,Bioinformatics.Bioinformatics,The MIT Press,1998
- [6] GenomeNet_WWW_server <http://www.genome.ad.jp/index.html>
- [7] J.F. クロー/安田徳一 訳, 基礎集団遺伝学, 培風館,1998
- [8] J.F. クロー/木村資生, 太田朋子 共訳, 遺伝学概説,1991
- [9] 木村資生 編, 分子進化学入門, 培風館,1984
- [10] 木村資生, 分子進化の中立説, 紀伊國屋書店,1986
- [11] ジョン・メイナード＝スミス/巖佐庸, 原田祐子 訳, 進化遺伝学,1995
- [12] 宮田 隆 編, 分子進化, 共立出版,1998
- [13] 長谷川政美, 岸野洋久, 分子系統学, 岩波書店,1996

- [14] 根井正利/五條堀孝, 斎藤成也 共訳, 分子進化遺伝学, 培風館,1990
- [15] 三中信宏, 生物系統学, 東京大学出版会,1997
- [16] 吉川寛監修/宮田隆, 五條堀孝 編, ゲノム情報を読む, 共立出版,1997
- [17] 金久實, ヒューマンゲノム計画, 共立出版,1997
- [18] 金久實, ゲノム情報への招待, 共立出版,1996
- [19] Subhash Sharma,Applied Multivariate Techniques,John Wiley & Sons,Inc.,1996
- [20] 大隈昇,L. ルバール,A. モリノウ,K.M. ワーウィック, 馬場康維. 記述的多変量解析法, 日科技連,1994
- [21] Bruce Alberts,Dennis Bray,Julian Lewis,Martin Raff,Keith Roberts,James D.Watson. /監訳, 中村桂子, 藤山秋佐夫, 松原謙一. 細胞の分子生物学第 3 版, 教育社、1995
- [22] Bruce Alberts,Dennis Bray,Alexander Johnson,Julian Lewis,Martin Raff,Keith Roberts,Peter Walter/監訳, 中村桂子, 藤山秋佐夫, 松原謙一.Essential 細胞生物学, 南江堂,1999
- [23] 石橋貞彦,ATP と代謝制御, 東京大学出版会,1989
- [24] 高木利久, 金久實 編, ゲノムネットのデータベース利用法 [第 2 版], 共立出版,1998

第 A 章

ENZYME データベースから得られた、 PATHWAY マップの詳細

EC 4.1.1.32	MAP00020	Citrate cycle (TCA cycle)
	MAP00620	Pyruvate metabolism
EC 2.3.1.61	MAP00020	Citrate cycle (TCA cycle)
	MAP00310	Lysine degradation
EC 4.1.3.6	MAP00020	Citrate cycle (TCA cycle)
	MAP00720	Reductive carboxylate cycle (CO ₂ fixation)
EC 4.1.3.7	MAP00020	Citrate cycle (TCA cycle)
	MAP00630	Glyoxylate and dicarboxylate metabolism
EC 4.1.3.8	MAP00020	Citrate cycle (TCA cycle)

EC 1.8.1.4	MAP00010	Glycolysis / Gluconeogenesis
	MAP00020	Citrate cycle (TCA cycle)
	MAP00260	"Glycine, serine and threonine metabolism"
	MAP00620	Pyruvate metabolism
EC 4.1.1.49	MAP00020	Citrate cycle (TCA cycle)
	MAP00620	Pyruvate metabolism
	MAP00710	Carbon fixation
EC 1.1.1.41	MAP00020	Citrate cycle (TCA cycle)
EC 1.3.99.1	MAP00020	Citrate cycle (TCA cycle)
	MAP00190	Oxidative phosphorylation
	MAP00650	Butanoate metabolism
	MAP00720	Reductive carboxylate cycle (CO2 fixation)
EC 4.1.3.34	MAP00020	Citrate cycle (TCA cycle)
EC 1.1.1.42	MAP00020	Citrate cycle (TCA cycle)
	MAP00480	Glutathione metabolism
	MAP00720	Reductive carboxylate cycle (CO2 fixation)
EC 1.1.1.37	MAP00020	Citrate cycle (TCA cycle)
	MAP00620	Pyruvate metabolism
	MAP00630	Glyoxylate and dicarboxylate metabolism
	MAP00710	Carbon fixation
	MAP00720	Reductive carboxylate cycle (CO2 fixation)
	MAP00020	Citrate cycle (TCA cycle)
EC 6.2.1.4	MAP00020	Citrate cycle (TCA cycle)
	MAP00640	Propanoate metabolism
EC 6.2.1.5	MAP00020	Citrate cycle (TCA cycle)
	MAP00640	Propanoate metabolism
	MAP00660	C5-Branched dibasic acid metabolism
	MAP00720	Reductive carboxylate cycle (CO2 fixation)
EC 1.2.4.2	MAP00020	Citrate cycle (TCA cycle)
	MAP00310	Lysine degradation
	MAP00380	Tryptophan metabolism
EC 6.4.1.1	MAP00020	Citrate cycle (TCA cycle)
	MAP00252	Alanine and aspartate metabolism
	MAP00620	Pyruvate metabolism
EC 1.2.7.3	MAP00020	Citrate cycle (TCA cycle)
	MAP00720	Reductive carboxylate cycle (CO2 fixation)
EC 2.8.3.10	MAP00020	Citrate cycle (TCA cycle)
EC 4.2.1.2	MAP00020	Citrate cycle (TCA cycle)
	MAP00720	Reductive carboxylate cycle (CO2 fixation)
EC 4.2.1.3	MAP00020	Citrate cycle (TCA cycle)
	MAP00630	Glyoxylate and dicarboxylate metabolism
	MAP00720	Reductive carboxylate cycle (CO2 fixation)

第 B 章

本論文で対象とした酵素データ

```
ENTRY      EC 4.1.1.32
NAME        Phosphoenolpyruvate carboxykinase (GTP)
            Phosphoenolpyruvate carboxylase
            Phosphopyruvate carboxylase
CLASS       Lyases
            Carbon-carbon lyases
            Carboxy-lyases
SYSNAME     GTP:oxaloacetate carboxy-lyase (transphosphorylating)
REACTION    GTP + Oxaloacetate = GDP + Phosphoenolpyruvate + CO2;
            ITP + Oxaloacetate = IDP + Phosphoenolpyruvate + CO2
SUBSTRATE   GTP
            Oxaloacetate
            ITP
PRODUCT     GDP
            Phosphoenolpyruvate
            CO2
COMMENT     ITP can act as phosphate donor.
PATHWAY     PATH: MAP00020 Citrate cycle (TCA cycle)
            PATH: MAP00620 Pyruvate metabolism
GENES       MTU: Rv0211(pckA)
            CTR: CT710
            CPN: CPn0851
            PHO: PH0312
            PAB: PAB1253
            CEL: R11A5.4 W05G11.6
            MMU: 97501(Pck1)
            HSA: 5105(PCK1) 5106(PCK2)
DISEASE     MIM: 261680 Phosphoenolpyruvate carboxykinase-1 (soluble);
            ?Hypoglycemia due to PCK1 deficiency (1)
MOTIF       PS: PS00505 F-P-S-A-C-G-K-T-N
DBLINKS     University of Geneva ENZYME DATA BANK: 4.1.1.32
            WIT (What Is There) Metabolic Reconstruction: 4.1.1.32
            BRENDA, the Enzyme Database: 4.1.1.32
```

ENTRY EC 2.3.1.61

NAME Dihydrolipoamide S-succinyltransferase

CLASS Transferases
 Acyltransferases
 Acyltransferases

SYSNAME Succinyl-CoA:dihydrolipoamide S-succinyltransferase

REACTION Succinyl-CoA + Dihydrolipoamide = CoA + S-Succinyldihydrolipoamide

SUBSTRATE Succinyl-CoA
 Dihydrolipoamide

PRODUCT CoA
 S-Succinyldihydrolipoamide

COMMENT A lipoyl-protein; component of the multienzyme 2-oxoglutarate dehydrogenase complex.

PATHWAY PATH: MAP00020 Citrate cycle (TCA cycle)
 PATH: MAP00310 Lysine degradation

GENES ECO: b0727(sucB)
 HIN: HI1661(sucB)
 RPR: RP179(sucB)
 BSU: odhB(odhB)
 MTU: Rv2215(sucB)
 CTR: CT055
 CPN: CPn0377
 SCE: YDR148C(KGD2)
 CEL: W02F12.5

DISEASE MIM: 126063 Dihydrolipoamide S-succinyltransferase (E2 component of

MOTIF PS: PS00189 [GN]-x(2)-[LIVF]-x(5)-[LIVFC]-x(2)-[LIVFA]-x(3)-K-[STAIV]-[STAVQDN]-x(2)-[LIVMFS]-x(5)-[GCN]-x-[LIVMFY]

STRUCTURES PDB: 1GHJ 1GHK 1HBR 1BAL 1BBL 1E20 1PMR

DBLINKS University of Geneva ENZYME DATA BANK: 2.3.1.61
 WIT (What Is There) Metabolic Reconstruction: 2.3.1.61
 BRENDA, the Enzyme Database: 2.3.1.61
 SCOP (Structural Classification of Proteins): 2.3.1.61

ENTRY EC 4.1.3.6

NAME Citrate lyase
Citrase
Citratase
Citritase
Citridesmolase
Citrate aldolase

CLASS Lyases
Carbon-carbon lyases
Oxo-acid-lyases

SYSNAME [Citrate oxaloacetate-lyase ((pro-3S)-CH₂COO- -->acetate)]

REACTION Citrate = Acetate + Oxaloacetate

SUBSTRATE Citrate

PRODUCT Acetate
Oxaloacetate

COMMENT The enzyme can be dissociated into components, two of which are identical with EC 2.8.3.10 and 4.1.3.34. EC 3.1.2.16 deacetylates and inactivates the enzyme.

PATHWAY PATH: MAP00020 Citrate cycle (TCA cycle)
PATH: MAP00720 Reductive carboxylate cycle (CO₂ fixation)

GENES ECO: b0615(citF) b0616(citE)
HIN: HI0022(citF) HI0023(citE)
MTU: Rv2498c(citE)
APE: APE0311
CEL: C01G10.7

DBLINKS University of Geneva ENZYME DATA BANK: 4.1.3.6
WIT (What Is There) Metabolic Reconstruction: 4.1.3.6
BRENDA, the Enzyme Database: 4.1.3.6

ENTRY EC 4.1.3.7

NAME Citrate (si)-synthase
Condensing enzyme
Citrate condensing enzyme
Citrogenase
Oxaloacetate transacetase

CLASS Lyases
Carbon-carbon lyases
Oxo-acid-lyases

SYSNAME Citrate oxaloacetate-lyase ((pro-3S)-CH₂COO- -> acetyl-CoA)

REACTION Citrate + CoA = Acetyl-CoA + H₂O + Oxaloacetate

SUBSTRATE Citrate
CoA

PRODUCT Acetyl-CoA
H₂O
Oxaloacetate

PATHWAY PATH: MAP00020 Citrate cycle (TCA cycle)
PATH: MAP00630 Glyoxylate and dicarboxylate metabolism

GENES EC0: b0720(gltA)
HPY: HP0026(gltA)
HPJ: jhp0022
RPR: RP844(gltA)
BSU: citA(citA) citZ(citZ) mmgD(mmgD)
MTU: Rv0889c(citA) Rv0896(gltA2) Rv1131(gltA1)
SYN: sll0401(gltA)
AAE: aq_150(gltA)
TMA: TMO290
MTH: MTH1726 MTH962
AFU: AF1340(citZ)
APE: APE1713
SCE: YCR005C(CIT2) YNR001C(CIT1) YPR001W(CIT3)
MMU: 88529
HSA: 1431

DISEASE MIM: 118950 Citrate synthase, mitochondrial

MOTIF PS: PS00342 [STAGCN]-[RKH]-[LIVMAFY]>
PS: PS00480 G-[FYA]-[GA]-H-x-[IV]-x(1,2)-[RKT]-x(2)-D-[PS]-R

STRUCTURES PDB: 1CSI 2CTS 1AJ8 1AL6 1AMZ 1CSC 1CSH 1CTS 1CSR 1CSS
2CSC 6CTS 3CSC 3CTS 4CSC 4CTS 5CSC 5CTS 6CSC

DBLINKS University of Geneva ENZYME DATA BANK: 4.1.3.7
WIT (What Is There) Metabolic Reconstruction: 4.1.3.7
BRENDA, the Enzyme Database: 4.1.3.7
SCOP (Structural Classification of Proteins): 4.1.3.7

ENTRY EC 4.1.3.8

NAME ATP citrate (pro-S)-lyase
Citrate cleavage enzyme

CLASS Lyases
Carbon-carbon lyases
Oxo-acid-lyases

SYSNAME ATP:citrate oxaloacetate-lyase ((pro-3S)-CH₂COO- ->acetyl-CoA)
(ATP-dephosphorylating)

REACTION ATP + Citrate + CoA = ADP + Orthophosphate + Acetyl-CoA +
Oxaloacetate

SUBSTRATE ATP
Citrate
CoA

PRODUCT ADP
Orthophosphate
Acetyl-CoA
Oxaloacetate

COMMENT The enzyme can be dissociated into components, two of which are
identical with EC 4.1.3.34 and 6.2.1.18.

PATHWAY PATH: MAP00020 Citrate cycle (TCA cycle)

GENES CEL: B0365.1 D1005.1
HSA: 47(ACLY)

DISEASE MIM: 108728 ATP citrate lyase

MOTIF PS: PS00399 G-x(2)-A-x(4,7)-[RQT]-[LIVMF]-G-H-[AS]-[GH]
PS: PS01216 S-[KR]-S-G-[GT]-[LIVM]-[GST]-x-[EQ]-x(8,10)-G-x(4)-
[LIVM]-[GA]-[LIVM]-G-G-D
PS: PS01217 G-x-[IV]-x(2)-[LIVMF]-x-[NA]-G-[GA]-G-[LA]-[STAV]-x(4)-
D-x-[LIVM]-x(3)-G-[GRE]

DBLINKS University of Geneva ENZYME DATA BANK: 4.1.3.8
WIT (What Is There) Metabolic Reconstruction: 4.1.3.8
BRENDA, the Enzyme Database: 4.1.3.8

ENTRY EC 1.8.1.4

NAME Dihydrolipoamide dehydrogenase
Diaphorase
Lipoamide reductase (NADH)
Lipoyl dehydrogenase

CLASS Oxidoreductases
Acting on a sulfur group of donors
With NAD+ or NADP+ as acceptor

SYSNAME Dihydrolipoamide:NAD+ oxidoreductase

REACTION Dihydrolipoamide + NAD+ = Lipoamide + NADH

SUBSTRATE Dihydrolipoamide
Methylene blue
NAD+

PRODUCT Lipoamide
NADH

COFACTOR FAD

COMMENT A flavoprotein (FAD). Component of the multienzyme pyruvate dehydrogenase complex and 2-oxoglutarate dehydrogenase complex. This enzyme was first shown to catalyse the oxidation of NADH by methylene blue; this activity was called diaphorase. Formerly EC 1.6.4.3.

PATHWAY PATH: MAP00010 Glycolysis / Gluconeogenesis
PATH: MAP00020 Citrate cycle (TCA cycle)
PATH: MAP00260 Glycine, serine and threonine metabolism
PATH: MAP00620 Pyruvate metabolism

GENES ECO: b0116(lpdA)
HIN: HI1231(lpdA)
RPR: RP460(pdhd) RP805(pdhd)
BSU: pdhd(pdhd) yqiV
MGE: MG271(pdhd)
MPN: F11_orf457(pdhd)
MTU: Rv0462 Rv0794c(lpdB) Rv3303c(lpdA)
CTR: CT557
CPN: CPn0833
SYN: slr1096(phdD)
AAE: aq_736(lpdA)
TMA: TMO381
MJA: MJ0636
MTH: MTH1648
APE: APE1669
SCE: YFL018C(LPD1)
CEL: LLC1.3
MMU: 107450(Dld)

HSA: 1738(DLD)
 DISEASE MIM: 246900 Dihydrolipoamide dehydrogenase (E3 component of
 pyruvate dehydrogenase; Lipoamide dehydrogenase
 deficiency (3)
 MOTIF PS: PS00076 G-G-x-C-[LIVA]-x(2)-G-C-[LIVM]-P
 STRUCTURES PDB: 1EBD 1LPF 1LVL 2PDD 2PDE 3LAD
 DBLINKS University of Geneva ENZYME DATA BANK: 1.8.1.4
 WIT (What Is There) Metabolic Reconstruction: 1.8.1.4
 BRENDA, the Enzyme Database: 1.8.1.4
 SCOP (Structural Classification of Proteins): 1.8.1.4

ENTRY EC 4.1.1.49

NAME Phosphoenolpyruvate carboxykinase (ATP)
Phosphopyruvate carboxylase (ATP)
Phosphoenolpyruvate carboxylase
Phosphoenolpyruvate carboxykinase

CLASS Lyases
Carbon-carbon lyases
Carboxy-lyases

SYSNAME ATP:oxaloacetate carboxy-lyase (transphosphorylating)

REACTION $\text{ATP} + \text{Oxaloacetate} = \text{ADP} + \text{Phosphoenolpyruvate} + \text{CO}_2$

SUBSTRATE ATP
Oxaloacetate

PRODUCT ADP
Phosphoenolpyruvate
CO₂

PATHWAY PATH: MAP00020 Citrate cycle (TCA cycle)
PATH: MAP00620 Pyruvate metabolism
PATH: MAP00710 Carbon fixation

GENES ECO: b3403(pckA)
HIN: HI0809(pckA)
BSU: pckA(pckA)
SCE: YKR097W(PCK1)

MOTIF PS: PS00532 L-I-G-D-D-E-H-x-W-x-[DE]-x-G-[IV]-x-N

STRUCTURES PDB: 1AQ2 1AYL 1OEN

DBLINKS University of Geneva ENZYME DATA BANK: 4.1.1.49
WIT (What Is There) Metabolic Reconstruction: 4.1.1.49
BRENDA, the Enzyme Database: 4.1.1.49
SCOP (Structural Classification of Proteins): 4.1.1.49

ENTRY EC 1.1.1.41

NAME Isocitrate dehydrogenase (NAD+)
Isocitric dehydrogenase
beta-Ketoglutaric-isocitric carboxylase

CLASS Oxidoreductases
Acting on the CH-OH group of donors
With NAD+ or NADP+ as acceptor

SYSNAME Isocitrate:NAD+ oxidoreductase(decarboxylating)

REACTION Isocitrate + NAD+ = 2-Oxoglutarate + CO2 + NADH

SUBSTRATE Isocitrate
NAD+
(1R,2S)-1-Hydroxypropane-1,2,3-tricarboxylate

PRODUCT 2-Oxoglutarate
CO2
NADH

COMMENT The isomer of isocitrate involved is (1R,2S)-1-hydroxypropane-
-1,2,3-tricarboxylate, formerly termed threo-Ds-isocitrate (4109).
Does not decarboxylate added oxalosuccinate.

PATHWAY PATH: MAP00020 Citrate cycle (TCA cycle)

GENES SCE: YNL037C(IDH1) YOR136W(IDH2)
CEL: C37E2.1 F35G12.2 F43G9.1
HSA: 3419(IDH3A) 3420(IDH3B) 3421(IDH3G)

DISEASE MIM: 300089 Isocitrate dehydrogenase 3 (NAD+), gamma
MIM: 601149 Isocitrate dehydrogenase 3 (NAD+) alpha

MOTIF PS: PS00470 [NS]-[LIMYT]-[FYDN]-G-[DNT]-[IMVY]-x-[STGDN]-[DN]-x(2)-
[SGAP]-x(3,4)-G-[STG]-[LIVMPA]-G-[LIVMF]

DBLINKS University of Geneva ENZYME DATA BANK: 1.1.1.41
WIT (What Is There) Metabolic Reconstruction: 1.1.1.41
BRENDA, the Enzyme Database: 1.1.1.41

ENTRY EC 1.3.99.1

NAME Succinate dehydrogenase
Succinic dehydrogenase
Fumarate reductase
Fumaric hydrogenase

CLASS Oxidoreductases
Acting on the CH-CH group of donors
With other acceptors

SYSNAME Succinate:(acceptor) oxidoreductase

REACTION Succinate + Acceptor = Fumarate + Reduced acceptor

SUBSTRATE Succinate
Acceptor

PRODUCT Fumarate
Reduced acceptor

COFACTOR FAD
Iron
Sulfur

COMMENT A flavoprotein (FAD) containing iron-sulfur centres. A component of EC 1.3.5.1. (EC 1.3.5.1 Succinate dehydrogenase)

PATHWAY PATH: MAP00020 Citrate cycle (TCA cycle)
PATH: MAP00190 Oxidative phosphorylation
PATH: MAP00650 Butanoate metabolism
PATH: MAP00720 Reductive carboxylate cycle (CO2 fixation)

GENES ECO: b0721(sdhC) b0722(sdhD) b0723(sdhA) b0724(sdhB) b4151(frdD)
b4152(frdC) b4153(frdB) b4154(frdA)
HIN: HI0832(frdD) HI0833(frdC) HI0834(frdB) HI0835(frdA)
HPY: HP0191(frdB) HP0192(frdA)
HPJ: jhp0177 jhp0178
RPR: RP044(sdhB) RP126(sdhC) RP127(sdhD) RP128(sdhA)
BSU: sdhA(sdhA) sdhB(sdhB) sdhC(sdhC)
MTU: Rv0247c Rv0248c Rv1552(frdA) Rv1553(frdB) Rv1554(frdC)
Rv1555(frdD) Rv3316(sdhC) Rv3317(sdhD) Rv3318(sdhA)
Rv3319(sdhB)
CTR: CT591 CT592 CT593
CPN: CPn0788 CPn0789 CPn0790
SYN: s110823(sdhB) s111625(sdhB) slr1233(frdA)
AAE: aq_553(frdB1) aq_594(frdA) aq_655(frdB2)
TMA: TM0427
MJA: MJ0033 MJ0092
MTH: MTH1502 MTH1850
AFU: AF0681(sdhA) AF0682(sdhB) AF0683(sdhC) AF0684(sdhD)
AF1463(fdrA)
APE: APE0946 APE0950

HSA: 6390(SDHB)

DISEASE MIM: 600857 Succinate dehydrogenase complex, subunit A,
flavoprotein (Fp); Leigh syndrome (3)

MOTIF PS: PS00190 C-{CPWHF}-{CPWR}-C-H-{CFYW}
PS: PS00197 C-{C}-{C}-[GA]-{C}-C-[GAST]-{CPDEKRHFYW}-C
PS: PS00198 C-x(2)-C-x(2)-C-x(3)-C-[PEG]
PS: PS00504 R-[ST]-H-[ST]-x(2)-A-x-G-G

DBLINKS University of Geneva ENZYME DATA BANK: 1.3.99.1
WIT (What Is There) Metabolic Reconstruction: 1.3.99.1
BRENDA, the Enzyme Database: 1.3.99.1

ENTRY EC 4.1.3.34

NAME Citryl-CoA lyase

CLASS Lyases
Carbon-carbon lyases
Oxo-acid-lyases

SYSNAME (3S)-Citryl-CoA oxaloacetate-lyase

REACTION (3S)-Citryl-CoA = Acetyl-CoA + Oxaloacetate

SUBSTRATE (3S)-Citryl-CoA

PRODUCT Acetyl-CoA
Oxaloacetate

COMMENT The enzyme is a component of EC 4.1.3.6 and 4.1.3.8. Also acts on (3S)-citryl thioacyl-carrier protein.

PATHWAY PATH: MAP00020 Citrate cycle (TCA cycle)

GENES HIN: HI0023(citE)

DBLINKS University of Geneva ENZYME DATA BANK: 4.1.3.34
WIT (What Is There) Metabolic Reconstruction: 4.1.3.34
BRENDA, the Enzyme Database: 4.1.3.34

ENTRY EC 1.1.1.42

NAME Isocitrate dehydrogenase (NADP+)
Oxalosuccinate decarboxylase

CLASS Oxidoreductases
Acting on the CH-OH group of donors
With NAD+ or NADP+ as acceptor

SYSNAME Isocitrate:NADP+ oxidoreductase (decarboxylating)

REACTION Isocitrate + NADP+ = 2-Oxoglutarate + CO2 + NADPH

SUBSTRATE Isocitrate
(1R,2S)-1-Hydroxypropane-1,2,3-tricarboxylate
NADP+

PRODUCT 2-Oxoglutarate
CO2
NADPH

COMMENT The isomer of isocitrate involved is (1R,2S)-1-hydroxypropane-1,2,3-tricarboxylate, formerly termed threo-Ds-isocitrate.
Also decarboxylates added oxalosuccinate.

PATHWAY PATH: MAP00020 Citrate cycle (TCA cycle)
PATH: MAP00480 Glutathione metabolism
PATH: MAP00720 Reductive carboxylate cycle (CO2 fixation)

GENES ECO: b1136(icd)
HPY: HP0027(icd)
HPJ: jhp0023
RPR: RP265(icd)
BSU: citC(citC)
MTU: Rv3339c(icd1)
SYN: slr1289(icd)
AAE: aq_1512(icd)
TMA: TM1148
MJA: MJ0720 MJ1596(icd)
MTH: MTH184
AFU: AF0647(icd)
APE: APE0689
SCE: YDL066W(IDP1) YLR174W(IDP2) YNL009W(IDP3)
CEL: C34F6.8 F59B8.2
MMU: 96413(Idh1)
HSA: 3417(IDH1) 3418(IDH2)

DISEASE MIM: 147650 Isocitrate dehydrogenase, mitochondrial

MOTIF PS: PS00470 [NS]-[LIMYT]-[FYDN]-G-[DNT]-[IMVY]-x-[STGDN]-[DN]-x(2)-[SGAP]-x(3,4)-G-[STG]-[LIVMPA]-G-[LIVMF]

STRUCTURES PDB: 8ICD 7ICD 6ICD 9ICD 5ICD 1CW1 1IDD 1IDC 4ICD 3ICD
1SJS 1ISO 1IKA 1IDF 1GRP 1AI2 1AI3 1BL5 1IDE 1CW4
1CW7 1GRO

DBLINKS University of Geneva ENZYME DATA BANK: 1.1.1.42
 WIT (What Is There) Metabolic Reconstruction: 1.1.1.42
 BRENDA, the Enzyme Database: 1.1.1.42
 SCOP (Structural Classification of Proteins): 1.1.1.42

ENTRY EC 1.1.1.37

NAME Malate dehydrogenase
Malic dehydrogenase

CLASS Oxidoreductases
Acting on the CH-OH group of donors
With NAD+ or NADP+ as acceptor

SYSNAME (S)-Malate:NAD+ oxidoreductase

REACTION (S)-Malate + NAD+ = Oxaloacetate + NADH

SUBSTRATE (S)-Malate
2-Hydroxydicarboxylic acid
NAD+

PRODUCT Oxaloacetate
NADH

COMMENT Also oxidizes some other 2-hydroxydicarboxylic acids.

PATHWAY PATH: MAP00020 Citrate cycle (TCA cycle)
PATH: MAP00620 Pyruvate metabolism
PATH: MAP00630 Glyoxylate and dicarboxylate metabolism
PATH: MAP00710 Carbon fixation
PATH: MAP00720 Reductive carboxylate cycle (CO2 fixation)

GENES ECO: b3236(mdh)
HIN: HI1210(mdh)
RPR: RP376(mdh)
BSU: citH(citH) yjmC
MTU: Rv1240(mdh)
CTR: CT376
CPN: CPn1028
SYN: sl10891(citH)
AAE: aq_1665(mdh2) aq_1782(mdh1)
MJA: MJ0490(ldh) MJ1425(mdh)
MTH: MTH1205 MTH188
AFU: AF0855(mdhA)
PHO: PH1277
PAB: PAB1791
APE: APE0672
SCE: YDL078C(MDH3) YKL085W(MDH1) YOL126C(MDH2)
CEL: F20H11.3
MMU: 97050(Mor1) 97051(Mor2)
HSA: 4190(MDH1) 4191(MDH2)

DISEASE MIM: 154100 Malate dehydrogenase, mitochondrial
MIM: 154200 Malate dehydrogenase, soluble

MOTIF PS: PS00068 [LIVM]-T-[TRKMN]-L-D-x(2)-R-[STA]-x(3)-[LIVMFY]
PS: PS00342 [STAGCN]-[RKH]-[LIVMAFY]>

STRUCTURES PDB: 1EMD 1BDM 1BMD 1CME 1HLP 1MLD 2CMD 4MDH 5MDH

DBLINKS University of Geneva ENZYME DATA BANK: 1.1.1.37
 WIT (What Is There) Metabolic Reconstruction: 1.1.1.37
 BRENDA, the Enzyme Database: 1.1.1.37
 SCOP (Structural Classification of Proteins): 1.1.1.37

ENTRY EC 6.2.1.4

NAME Succinate--CoA ligase (GDP-forming)
Succinyl-CoA synthetase (GDP-forming)

CLASS Ligases
Forming carbon-sulfur bonds
Acid--thiol ligases

SYSNAME Succinate:CoA ligase (GDP-forming)

REACTION GTP + Succinate + CoA = GDP + Orthophosphate + Succinyl-CoA;
ITP + Succinate + CoA = IDP + Orthophosphate + Succinyl-CoA;
GTP + Itaconate + CoA = GDP + Orthophosphate + Itaconyl-CoA;
ITP + Itaconate + CoA = IDP + Orthophosphate + Itaconyl-CoA

SUBSTRATE GTP
Succinate
CoA
Itaconate
ITP

PRODUCT GDP
Orthophosphate
Succinyl-CoA
Itaconyl-CoA
IDP

COMMENT Itaconate can act instead of succinate and ITP instead of GTP.

PATHWAY PATH: MAP00020 Citrate cycle (TCA cycle)
PATH: MAP00640 Propanoate metabolism

GENES SCE: YGR244C(LSC2) YOR142W(LSC1)
CEL: C05G5.4 C50F7.4 F47B10.1

MOTIF PS: PS00399 G-x(2)-A-x(4,7)-[RQT]-[LIVMF]-G-H-[AS]-[GH]
PS: PS01216 S-[KR]-S-G-[GT]-[LIVM]-[GST]-x-[EQ]-x(8,10)-G-x(4)-
[LIVM]-[GA]-[LIVM]-G-G-D
PS: PS01217 G-x-[IV]-x(2)-[LIVMF]-x-[NA]-G-[GA]-G-[LA]-[STAV]-x(4)-
D-x-[LIVM]-x(3)-G-[GREA]

DBLINKS University of Geneva ENZYME DATA BANK: 6.2.1.4
WIT (What Is There) Metabolic Reconstruction: 6.2.1.4
BRENDA, the Enzyme Database: 6.2.1.4

ENTRY EC 6.2.1.5

NAME Succinate--CoA ligase (ADP-forming)
Succinyl-CoA synthetase (ADP-forming)

CLASS Ligases
Forming carbon-sulfur bonds
Acid--thiol ligases

SYSNAME Succinate:CoA ligase (ADP-forming)

REACTION $\text{ATP} + \text{Succinate} + \text{CoA} = \text{ADP} + \text{Orthophosphate} + \text{Succinyl-CoA}$

SUBSTRATE ATP
Succinate
CoA

PRODUCT ADP
Orthophosphate
Succinyl-CoA

PATHWAY PATH: MAP00020 Citrate cycle (TCA cycle)
PATH: MAP00640 Propanoate metabolism
PATH: MAP00660 C5-Branched dibasic acid metabolism
PATH: MAP00720 Reductive carboxylate cycle (CO2 fixation)

GENES ECO: b0728(sucC) b0729(sucD)
HIN: HI1196(sucC) HI1197(sucD)
RPR: RP432(sucD) RP433(sucC)
BSU: sucC(sucC) sucD(sucD)
MTU: Rv0951(sucC) Rv0952(sucD)
CTR: CT821 CT822
CPN: CPn0973 CPn0974
SYN: sll1023(sucC) sll1557(sucD)
AAE: aq_1306(sucC1) aq_1620(sucC2) aq_1622(sucD2) aq_1888(sucD1)
MJA: MJ0210(sucC) MJ1246(sucD)
MTH: MTH1036 MTH563
AFU: AF1539(sucD-1) AF1540(sucC-1) AF2185(sucD-2) AF2186(sucC-2)
APE: APE1065 APE1072
CEL: F23H11.3

MOTIF PS: PS00399 G-x(2)-A-x(4,7)-[RQT]-[LIVMF]-G-H-[AS]-[GH]
PS: PS01216 S-[KR]-S-G-[GT]-[LIVM]-[GST]-x-[EQ]-x(8,10)-G-x(4)-
[LIVM]-[GA]-[LIVM]-G-G-D
PS: PS01217 G-x-[IV]-x(2)-[LIVMF]-x-[NA]-G-[GA]-G-[LA]-[STAV]-x(4)-
D-x-[LIVM]-x(3)-G-[GREAA]

STRUCTURES PDB: 1SCU 2SCU

DBLINKS University of Geneva ENZYME DATA BANK: 6.2.1.5
WIT (What Is There) Metabolic Reconstruction: 6.2.1.5
BRENDA, the Enzyme Database: 6.2.1.5
SCOP (Structural Classification of Proteins): 6.2.1.5

ENTRY EC 1.2.4.2

NAME Oxoglutarate dehydrogenase (lipoamide)
Oxoglutarate decarboxylase
alpha-Ketoglutaric dehydrogenase

CLASS Oxidoreductases
Acting on the aldehyde or oxo group of donors
With a disulfide as acceptor

SYSNAME 2-Oxoglutarate:lipoamide 2-oxidoreductase(decarboxylating and acceptor-succinylating)

REACTION 2-Oxoglutarate + Lipoamide = S-Succinyldihydrolipoamide + CO₂

SUBSTRATE 2-Oxoglutarate
Lipoamide

PRODUCT S-Succinyldihydrolipoamide
CO₂

COFACTOR Thiamine diphosphate

COMMENT Requires thiamine diphosphate; component of the multienzyme 2-oxoglutarate dehydrogenase complex.

PATHWAY PATH: MAP00020 Citrate cycle (TCA cycle)
PATH: MAP00310 Lysine degradation
PATH: MAP00380 Tryptophan metabolism

GENES ECO: b0726(sucA)
HIN: HI1662(sucA)
RPR: RP180(sucA)
BSU: odhA(odhA)
MTU: Rv1248c(sucA)
CTR: CT054
CPN: CPn0378
SCE: YIL125W(KGD1)
CEL: T22B11.5 ZK836.2
HSA: 4967(OGDH)

DISEASE MIM: 203740 Oxoglutarate dehydrogenase (lipoamide);
Alpha-ketoglutarate dehydrogenase deficiency (1)

DBLINKS University of Geneva ENZYME DATA BANK: 1.2.4.2
WIT (What Is There) Metabolic Reconstruction: 1.2.4.2
BRENDA, the Enzyme Database: 1.2.4.2

ENTRY	EC 6.4.1.1
NAME	Pyruvate carboxylase Pyruvic carboxylase
CLASS	Ligases Forming carbon-carbon bonds
SYSNAME	Pyruvate:carbon-dioxide ligase (ADP-forming)
REACTION	ATP + Pyruvate + HC03- = ADP + Orthophosphate + Oxaloacetate
SUBSTRATE	ATP Pyruvate HC03-
PRODUCT	ADP Orthophosphate Oxaloacetate
COFACTOR	Biotin Manganese Zinc Acetyl-CoA
COMMENT	A biotinyl-protein containing manganese (animal tissues) or zinc (yeast). The animal enzyme requires acetyl-CoA.
PATHWAY	PATH: MAP00020 Citrate cycle (TCA cycle) PATH: MAP00252 Alanine and aspartate metabolism PATH: MAP00620 Pyruvate metabolism
GENES	BSU: pycA(pycA) MTU: Rv2967c(pca) AAE: aq_1517(pycB) SCE: YBR218C(PYC2) YGL062W(PYC1) CEL: D2023.2 MMU: 97520(Pcx) HSA: 5091
DISEASE	MIM: 266150 Pyruvate carboxylase; Pyruvate carboxylase deficiency (3)
MOTIF	PS: PS00188 [GN]-[DEQTR]-x-[LIVMFY]-x(2)-[LIVM]-x-[AIV]-M-K-[LMAT]-x(3)-[LIVM]-x-[SAV] PS: PS00866 [FYV]-[PS]-[LIVMC]-[LIVMA]-[LIVM]-[KR]-[PSA]-[STA]-x(3)-[SG]-G-x-[AG] PS: PS00867 [LIVMF]-[LIMN]-E-[LIVMCA]-N-[PATLIVM]-[KR]-[LIVMSTAC]
DBLINKS	University of Geneva ENZYME DATA BANK: 6.4.1.1 WIT (What Is There) Metabolic Reconstruction: 6.4.1.1 BRENDA, the Enzyme Database: 6.4.1.1

ENTRY	EC 3.1.2.3
NAME	Succinyl-CoA hydrolase Succinyl-CoA acylase
CLASS	Hydrolases Acting on ester bonds Thiolester hydrolases
SYSNAME	Succinyl-CoA hydrolase
REACTION	Succinyl-CoA + H2O = CoA + Succinate
SUBSTRATE	Succinyl-CoA H2O
PRODUCT	CoA Succinate
PATHWAY	PATH: MAP00020 Citrate cycle (TCA cycle)
DBLINKS	University of Geneva ENZYME DATA BANK: 3.1.2.3 WIT (What Is There) Metabolic Reconstruction: 3.1.2.3 BRENDA, the Enzyme Database: 3.1.2.3

ENTRY	EC 6.2.1.18
NAME	Citrate--CoA ligase
CLASS	Ligases
	Forming carbon-sulfur bonds
	Acid--thiol ligases
SYSNAME	Citrate:CoA ligase (ADP-forming)
REACTION	ATP + Citrate + CoA = ADP + Orthophosphate + (3S)-Citryl-CoA
SUBSTRATE	ATP
	Citrate
	CoA
PRODUCT	ADP
	Orthophosphate
	(3S)-Citryl-CoA
COMMENT	The enzyme is a component of EC 4.1.3.8.
PATHWAY	PATH: MAP00020 Citrate cycle (TCA cycle)
DBLINKS	University of Geneva ENZYME DATA BANK: 6.2.1.18
	WIT (What Is There) Metabolic Reconstruction: 6.2.1.18
	BRENDA, the Enzyme Database: 6.2.1.18

ENTRY EC 1.2.7.3

NAME 2-Oxoglutarate synthase

CLASS Oxidoreductases
Acting on the aldehyde or oxo group of donors
With an iron-sulfur protein as acceptor

SYSNAME 2-Oxoglutarate:ferredoxin 2-oxidoreductase (CoA-succinylating)

REACTION 2-Oxoglutarate + CoA + Oxidized ferredoxin = Succinyl-CoA +
CO₂ + Reduced ferredoxin

SUBSTRATE 2-Oxoglutarate
CoA
Oxidized ferredoxin

PRODUCT Succinyl-CoA
CO₂
Reduced ferredoxin

PATHWAY PATH: MAP00020 Citrate cycle (TCA cycle)
PATH: MAP00720 Reductive carboxylate cycle (CO₂ fixation)

GENES HPY: HP0588 HP0589 HP0590 HP0591
MTH: MTH1032 MTH1033 MTH1034 MTH1035
AFU: AF0468(korB) AF0469(korA) AF0470(korD) AF0471(korG)

DBLINKS University of Geneva ENZYME DATA BANK: 1.2.7.3
WIT (What Is There) Metabolic Reconstruction: 1.2.7.3
BRENDA, the Enzyme Database: 1.2.7.3

ENTRY EC 2.8.3.10

NAME Citrate CoA-transferase

CLASS Transferases
Transferring sulfur-containing groups
CoA-transferases

SYSNAME Acetyl-CoA: citrate CoA-transferase

REACTION Acetyl-CoA + Citrate = Acetate + (3S)-Citryl-CoA

SUBSTRATE Acetyl-CoA
Citrate

PRODUCT Acetate
(3S)-Citryl-CoA

COMMENT The enzyme is a component of EC 4.1.3.6. Also catalyses the transfer of thioacyl carrier protein from its acetyl thioester to citrate.

PATHWAY PATH: MAP00020 Citrate cycle (TCA cycle)

GENES HIN: HI0022(citF)

DBLINKS University of Geneva ENZYME DATA BANK: 2.8.3.10
WIT (What Is There) Metabolic Reconstruction: 2.8.3.10
BRENDA, the Enzyme Database: 2.8.3.10

ENTRY EC 4.2.1.2

NAME Fumarate hydratase
Fumarase

CLASS Lyases
Carbon-oxygen lyases
Hydro-lyases

SYSNAME (S)-Malate hydro-lyase

REACTION (S)-Malate = Fumarate + H2O

SUBSTRATE (S)-Malate

PRODUCT Fumarate
H2O

PATHWAY PATH: MAP00020 Citrate cycle (TCA cycle)
PATH: MAP00720 Reductive carboxylate cycle (CO2 fixation)

GENES ECO: b1611(fumC) b1612(fumA) b4122(fumB)
HIN: HI1398(fumC)
HPY: HP1325(fumC)
HPJ: jhp1245
RPR: RP665(fumC)
BSU: citG(citG)
MTU: Rv1098c(fum)
CTR: CT855
CPN: CPn1013
SYN: slr0018(fumC)
AAE: aq_1679(fumX) aq_1780(fumB)
TMA: TM0540 TM0541
MJA: MJ0617 MJ1294
MTH: MTH1115 MTH1735 MTH1910 MTH963
AFU: AF1098(fum-1) AF1099(fum-2)
PHO: PH1683 PH1684
PAB: PAB2030(fum-1) PAB2031(fum-2)
APE: APE1816
SCE: YPL262W(FUM1)
CEL: H14A12.2
HSA: 2271

DISEASE MIM: 136850 Fumarate hydratase; Fumarase deficiency (3)

MOTIF PS: PS00163 G-S-x(2)-M-x(2)-K-x-N

STRUCTURES PDB: 2FUS 1YFM 1FUR 1FUQ 1FUP 1FUO

DBLINKS University of Geneva ENZYME DATA BANK: 4.2.1.2
WIT (What Is There) Metabolic Reconstruction: 4.2.1.2
BRENDA, the Enzyme Database: 4.2.1.2
SCOP (Structural Classification of Proteins): 4.2.1.2

ENTRY EC 4.2.1.3

NAME Aconitate hydratase
Aconitase

CLASS Lyases
Carbon-oxygen lyases
Hydro-lyases

SYSNAME Citrate(isocitrate) hydro-lyase

REACTION Citrate = Isocitrate

SUBSTRATE Citrate
cis-Aconitate
H2O
Isocitrate

PRODUCT cis-Aconitate
H2O

COFACTOR Iron
Sulfur

COMMENT An iron-sulfur protein. cis-Aconitate is used to designate the isomer (Z)-prop-1-ene-1,2,3-tricarboxylate. Also converts isocitrate into cis-aconitate.
(Proc.Natl.Acad.Sci.USA (1991) 88, 10109-10113)

PATHWAY PATH: MAP00020 Citrate cycle (TCA cycle)
PATH: MAP00630 Glyoxylate and dicarboxylate metabolism
PATH: MAP00720 Reductive carboxylate cycle (CO2 fixation)

GENES ECO: b0118(acnB) b1276(acnA)
HPY: HP0779(acnB)
HPJ: jhp0716
RPR: RP799(acnA)
BSU: citB(citB)
MTU: Rv1475c(acn)
SYN: slr0665(acnB)
AAE: aq_1784(aco)
APE: APE1618
SCE: YJL200C YLR304C(AC01)
CEL: F54H12.1 ZK455.1
MMU: 87879(Aco1)
HSA: 48(AC01) 50(AC02)

DISEASE MIM: 100850 Aconitase, mitochondrial

MOTIF PS: PS00450 [LIVM]-x(2)-[GSACIVM]-x-[LIV]-[GTIV]-[STP]-C-x(0,1)-T-N-[GSTANI]-x(4)-[LIVMA]
PS: PS01244 G-x(2)-[LIVWPQ]-x(3)-[GAC]-C-[GSTAM]-[LIMPTA]-C-[LIMV]-[GA]

STRUCTURES PDB: 8ACN 1ACO 1BOJ 1BOK 1BOM 1C96 1C97 7ACN 6ACN 5ACN
1NIT 1AMJ 1AMI 1NIS 1FGH

DBLINKS University of Geneva ENZYME DATA BANK: 4.2.1.3
 WIT (What Is There) Metabolic Reconstruction: 4.2.1.3
 BRENDA, the Enzyme Database: 4.2.1.3
 SCOP (Structural Classification of Proteins): 4.2.1.3