International Journal of Bioinformatics and Biological Science: v.1 n.1 p.63-70. March, 2013

Identification and analysis of metabolic pathways in *A. thaliana* & *E. coli*

Prashant Ankur Jain[*1], A. K. Gupta[2], Krishna Misra[3] and Satendra Singh[1]

[1]Department of Computational Biology & Bioinformatics, JSBB, SHIATS, Allahabad-211007 ,India
[2]Department of Forensic Science, SHIATS, Allahabad-211007,India
[3]Division of Bioinformatics, IIITA, Allahabad-211007 ,India

[*]Corresponding Author: Prashant Ankur Jain prash1721@gmail.com

Received: 24 December 2012; Accepted: 03 January 2013

ABSTRACT

The prokaryotes and eukaryotes are separated by roughly three billion years of evolution, when bacterial and eukaryote division is thought to have taken place. During this time there have been countless chances for the genes in two organisms to diverge by mutation, to change gene structure by gene fusion or fission and to acquire new genes for a function by horizontal transfer or functional displacement of one gene by another within a genome. In the present study we have compared & analyzed the metabolic pathways of *A. thaliana* & *E.coli* to understand the extant of conservation of enzymes in their metabolic pathways. Present study indicate that the *A. thaliana*, a flowering plant and *E.col*, a gram negative bacteria were found sharing significant similarity in their metabolic pathways.

Keywords*:* A. thaliana, E.coli, *metabolic pathway*

BACKGROUND

Cellular metabolism represents a collection of enzymatic reactions and transport processes that convert metabolites into molecules capable of supporting cellular life. There is considerable interest in the processes underlying the evolution of cellular metabolism(Alves *et al*,2002). The existence of a core ensemble of metabolic reactions common to most organisms suggests that the global metabolic structure has been the subject of strong evolutionary constraint. Similarly, network connectivity properties suggest modular components typical of evolved systems and emergence of hub metabolites involved in many reactions by enzyme specialization(Brocks *et al* 1999). How metabolic networks function and change as organisms increased in complexity remains an important question, making metabolism an interesting model for the evolution of biomolecular networks. There is considerable evidence supporting the patchwork recruitment scenario. These patterns of structural homology appeared to be pervasive when structural assignments and sequence comparisons were used to analyze the small-molecule metabolism in *Escherichia coli*. Recruitment occurred with little regularity in these instances(Brown and Doolittle,1997). Such a

comparison may be much less successful in any pair of organisms due to lack of knowledge of their enzymes and pathways. So we have to select such model organisms which can be subjected to very extensive experimental characterization of their genes and proteins, including the determination of their complete genome sequence.

Regardless of the origin of the enzymes, during this time there have been countless chances for orthologous genes in the two organisms to diverge by mutation, to undergo recombination resulting in domain loss or accretion, and to change gene structure by gene fusion or fission. New genes for an existing function could be acquired by horizontal transfer or functional displacement of one gene by another within a genome(Lazcano and Miler 1996) In addition, many new genes have arisen by duplication and divergence to produce new enzymatic functions and pathways.

The extensive information available on the enzymes and pathways of small molecule metabolism in two different organisms allows us to determine the extent to which different evolutionary processes have taken place since they separated from their last common ancestor(Lawrence and Roth,1996) At present such a comparison would be much less successful in any other pair of organisms due to the lack of knowledge of their enzymes and pathways.

MATERIALS AND METHODS

The small molecule metabolic pathways and enzymes in *A. thaliana and* E. *coli* were modified from those present in the KEGG database. The format of the KEGG database makes it easy to draw parallels between organisms, but the reason we did not simply adopt the precise set of KEGG pathways and proteins is that the reconstruction of pathways in KEGG takes place purely on the basis of Enzyme Classification (EC) numbers. KEGG pathway database was used as a source of metabolic pathway information(Singh *et al* ,2012) . Metabolic pathway identification names and numbers of *A. thaliana* & *E.coli* were extracted from the KEGG database. Common and unique metabolic pathways of *A. thaliana* & *E.coli* were identified manually from the KEGG database (http://www.genome.jp/kegg/pathway.html). The corresponding protein sequences were retrieved from the KEGG database(La Kanehisa *et al*,2002).

RESULTS AND DISCUSSION

After comparison of the metabolic pathways in *A. Thaliana* and *E. coli* it was found that total 60 metabolic pathways were identified as common (Table 1). It showed that common metabolic pathways were playing key role in metabolic activities of two organisms. It was also observed that in most of the metabolic pathways, number of enzymes present in *A. thaliana* is much more in comparison with *E.coli*. While there are some metabolic pathways like Glyoxylate and dicarboxylate metabolism, Butanoate metabolism, Thiamine metabolism, Riboflavin metabolism, Nicotinate and nicotinamide metabolism and Biotin metabolism

Table 1: Common metabolic pathways identified in *A. Thaliana* and *E.coli*

S.No.	*A. thaliana*	Metabolic pathway	*E.coli*
1.	ath00010	Glycolysis / Gluconeogenesis	eco00010
2.	ath00020	Citrate cycle (TCA cycle)	eco00020
3.	ath00030	Pentose phosphate pathway	eco00030
4.	ath00040	Pentose and glucuronate interconversions	eco00040
5.	ath00051	Fructose and mannose metabolism	eco00051
6.	ath00052	Galactose metabolism	eco00052
7.	ath00053	Ascorbate and aldarate metabolism	eco00053
8.	ath00061	Fatty acid biosynthesis	eco00061
9.	ath00071	Fatty acid metabolism	eco00071
10.	ath00130	Ubiquinone and other terpenoid-quinone biosynthesis	eco00130
11.	ath00190	Oxidative phosphorylation	eco00190
12.	ath00230	Purine metabolism	eco00230
13.	ath00240	Pyrimidine metabolism	eco00240
14.	ath00250	Alanine, aspartate and glutamate metabolism	eco00250
15.	ath00260	Glycine, serine and threonine metabolism	eco00260
16.	ath00270	Cysteine and methionine metabolism	eco00270
17.	ath00280	Valine, leucine and isoleucine degradation	eco00280
18.	ath00290	Valine, leucine and isoleucine biosynthesis	eco00290
19.	ath00300	Lysine biosynthesis	eco00300
20.	ath00310	Lysine degradation	eco00310
21.	ath00330	Arginine and proline metabolism	eco00330
22.	ath00340	Histidine metabolism	eco00340
23.	ath00350	Tyrosine metabolism	eco00350
24.	ath00360	Phenylalanine metabolism	eco00360
25.	ath00380	Tryptophan metabolism	eco00380
26.	ath00400	Phenylalanine, tyrosine and tryptophan biosynthesis	eco00400
27.	ath00410	beta-Alanine metabolism	eco00410
28.	ath00430	Taurine and hypotaurine metabolism	eco00430
29.	ath00450	Selenoamino acid metabolism	eco00450
30.	ath00460	Cyanoamino acid metabolism	eco00460
31.	ath00480	Glutathione metabolism	eco00480
32.	ath00500	Starch and sucrose metabolism	eco00500
33.	ath00511	Other glycan degradation	eco00511
34.	ath00520	Amino sugar and nucleotide sugar metabolism	eco00520
35.	ath00561	Glycerolipid metabolism	eco00561
36.	ath00562	Inositol phosphate metabolism	eco00562
37.	ath00564	Glycerophospholipid metabolism	eco00564
38.	ath00590	Arachidonic acid metabolism	eco00590
39.	ath00592	alpha-Linolenic acid metabolism	eco00592
40.	ath00600	Sphingolipid metabolism	eco00600
41.	ath00620	Pyruvate metabolism	eco00620
42.	ath00630	Glyoxylate and dicarboxylate metabolism	eco00630

Contd.

S.No.	*A. thaliana*	Metabolic pathway	*E.coli*
43.	ath00640	Propanoate metabolism	eco00640
44.	ath00650	Butanoate metabolism	eco00650
45.	ath00660	C5-Branched dibasic acid metabolism	eco00660
46.	ath00670	One carbon pool by folate	eco00670
47.	ath00730	Thiamine metabolism	eco00730
48.	ath00740	Riboflavin metabolism	eco00740
49.	ath00750	Vitamin B6 metabolism	eco00750
50.	ath00760	Nicotinate and nicotinamide metabolism	eco00760
51.	ath00770	Pantothenate and CoA biosynthesis	eco00770
52.	ath00780	Biotin metabolism	eco00780
53.	ath00785	Lipoic acid metabolism	eco00785
54.	ath00790	Folate biosynthesis	eco00790
55.	ath00860	Porphyrin and chlorophyll metabolism	eco00860
56.	ath00900	Terpenoid backbone biosynthesis	eco00900
57.	ath00903	Limonene and pinene degradation	eco00903
58.	ath00910	Nitrogen metabolism	eco00910
59.	ath00920	Sulfur metabolism	eco00920
60.	ath01040	Biosynthesis of unsaturated fatty acids	eco01040

Table 2: Unique Metabolic pathways in *A. thaliana*

S.No	*A. thaliana*	Metabolic pathway
1.	ath00062	Fatty acid elongation in mitochondria
2.	ath00072	Synthesis and degradation of ketone bodies
3.	ath00100	Steroid biosynthesis
4.	ath00195	Photosynthesis
5.	ath00196	Photosynthesis - antenna proteins
6.	ath00232	Caffeine metabolism
7.	ath00510	N-Glycan biosynthesis
8.	ath00514	Other types of O-glycan biosynthesis
9.	ath00531	Glycosaminoglycan degradation
10.	ath00563	Glycosylphosphatidylinositol(GPI)-anchor biosynthesis
11.	ath00565	Ether lipid metabolism
12.	ath00603	Glycosphingolipid biosynthesis - globo series
13.	ath00604	Glycosphingolipid biosynthesis - ganglio series
14.	ath00710	Carbon fixation in photosynthetic organisms
15.	ath00901	Indole alkaloid biosynthesis
16.	ath00902	Monoterpenoid biosynthesis
17.	ath00904	Diterpenoid biosynthesis
18.	ath00905	Brassinosteroid biosynthesis
19.	ath00906	Glucosinolate biosynthesis

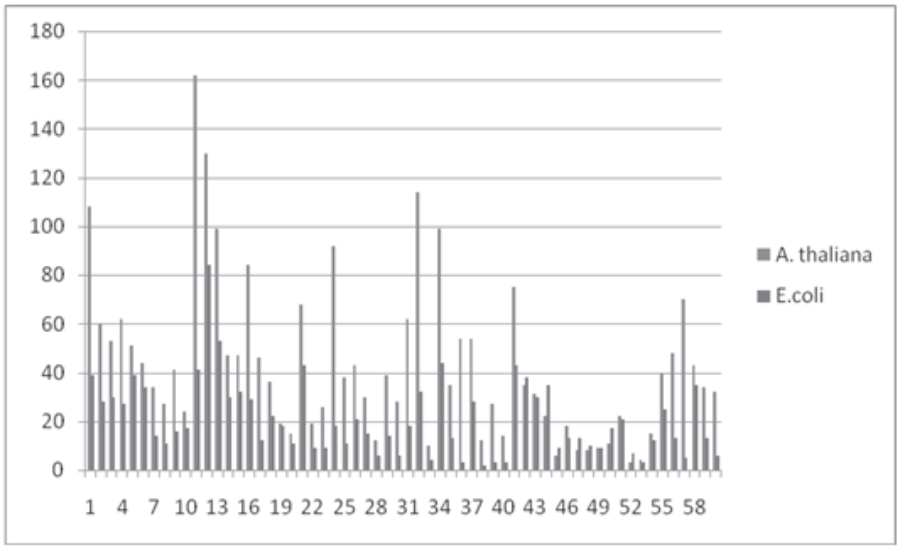

Figure 1: Total number of Enzymes present in common metabolic pathways of *A. thaliana* and *E.coli*

wherethe number of enzymes present is more in E.coli in comparison with *A. thaliana* (Table 4 /Figure 1). Also there are some metabolic pathways where the number of enzymes were almost equal for example Lysine biosynthesis, Propanoate metabolism, Vitamin B6 metabolism and Lipoic acid metabolism(Makarova *et al* ,1999). It was also found that *A. Thaliana* has total 88 metabolic pathways out of which 28 are unique with respect to *E.coli* (Table 2) whereas *E.coli* has total 84 metabolic pathways out of which 24 are unique with respect to *A. Thaliana* (Table 3). The unique enzymes in the two organisms were playing their role in other biological activities of the organism(Mojzsis *et al*,1996).

Until now, investigations of these evolutionary processes have been limited to individual instances or small set of occurrences, mostly identified by sequence comparison methods. Here we investigate, and to some extent quantify the frequency of these processes in a complete set of pathways between two distantly related organisms(Tsoka and Ouzounis ,2001). The large amount of information available about the pathways, functions and structures of enzymes in these organisms allow

Table 3: Unique Metabolic pathways in *E.coli*

S.No	*E.coli*	Metabolic Pathway
1.	eco00361	Chlorocyclohexane and chlorobenzene degradation
2.	eco00362	Benzoate degradation
3.	eco00363	Bisphenol degradation
4.	eco00364	Fluorobenzoate degradation
5.	eco00401	Novobiocin biosynthesis
6.	eco00440	Phosphonate and phosphinate metabolism
7.	eco00471	D-Glutamine and D-glutamate metabolism
8.	eco00473	D-Alanine metabolism
9.	eco00521	Streptomycin biosynthesis
10.	eco00523	Polyketide sugar unit biosynthesis
11.	eco00540	Lipopolysaccharide biosynthesis
12.	eco00550	Peptidoglycan biosynthesis
13.	eco00621	Dioxin degradation
14.	eco00622	Xylene degradation
15.	eco00623	Toluene degradation
16.	eco00625	Chloroalkane and chloroalkene degradation
17.	eco00626	Naphthalene degradation
18.	eco00627	Aminobenzoate degradation
19.	eco00633	Nitrotoluene degradation
20.	eco00642	Ethylbenzene degradation
21.	eco00680	Methane metabolism
22.	eco00720	Carbon fixation pathways in prokaryotes
23.	eco00930	Caprolactam degradation
24	eco01053	Biosynthesis of siderophore group nonribosomal peptides

Table 4: Total number of Enzymes present in common metabolic pathways of *A. thaliana* and *E.coli*

S.No.		*A. thaliana*	*E.coli*
1	Glycolysis/Gluconeogenesis	108	39
2	Citrate cycle (TCA cycle)	60	28
3	Pentose phosphate pathway	53	30
4	Pentose and glucuronate interconversions	62	27
5	Fructose and mannose metabolism	51	39
6	Galactose metabolism	44	34
7	Ascorbate and aldarate metabolism	34	14
8	Fatty acid biosynthesis	27	11
9	Fatty acid metabolism	41	16
10	Ubiquinone and other terpenoid-quinone biosynthesis	24	17
11	Oxidative phosphorylation	162	41
12	Purine metabolism1	30	84
13	Pyrimidine metabolism	99	53
14	Alanine, aspartate and glutamate metabolism	47	30

Contd.

S.No.		*A. thaliana*	*E.coli*
15	Glycine, serine and threonine metabolism	47	32
16	Cysteine and methionine metabolism	84	29
17	Valine, leucine and isoleucine degradation	46	12
18	Valine, leucine and isoleucine biosynthesis	36	22
19	Lysine biosynthesis	19	18
20	Lysine degradation	15	11
21	Arginine and proline metabolism	68	43
22	Histidine metabolism	19	9
23	Tyrosine metabolism	26	9
24	Phenylalanine metabolism	92	18
25	Tryptophan metabolism	38	11
26	Phenylalanine, tyrosine and tryptophan biosynthesis	43	21
27	beta-Alanine metabolism	30	15
28	Taurine and hypotaurine metabolism	12	6
29	Selenoamino acid metabolism	39	14
30	Cyanoamino acid metabolism	28	6
31	Glutathione metabolism	62	18
32	Starch and sucrose metabolism	114	32
33	Other glycan degradation	10	4
34	Amino sugar and nucleotide sugar metabolism	99	44
35	Glycerolipid metabolism	35	13
36	Inositol phosphate metabolism	54	3
37	Glycerophospholipid metabolism	54	28
38	Arachidonic acid metabolism	12	2
39	alpha-Linolenic acid metabolism	27	3
40	Sphingolipid metabolism	14	3
41	Pyruvate metabolism	75	43
42	Glyoxylate and dicarboxylate metabolism	35	38
43	\Propanoate metabolism	31	30
44	Butanoate metabolism	22	35
45	C5-Branched dibasic acid metabolism	6	9
46	One carbon pool by folate	18	13
47	Thiamine metabolism	8	13
48	Riboflavin metabolism	8	10
49	Vitamin B6 metabolism	9	9
50	Nicotinate and nicotinamide metabolism	11	17
51	Pantothenate and CoA biosynthesis	22	21
52	Biotin metabolism	3	7
53	Lipoic acid metabolism	4	3
54	Folate biosynthesis	15	12
55	Porphyrin and chlorophyll metabolism	40	25
56	Terpenoid backbone biosynthesis	48	13
57	Limonene and pinene degradation	70	5
58	Nitrogen metabolism	43	35
59	Sulfur metabolism	34	13
60	Biosynthesis of unsaturated fatty acids	32	6

us to study the evolutionary processes in small molecule metabolism(Wang *et al,*1999). These observations suggest the retrograde evolution model played a small part in the process of metabolic enzyme evolution.

CONCLUSION

Based on this study it can be concluded that the *A. Thaliana* and *E.coli,* from different kingdoms, have significant number of common metabolic pathways. But the number of enzymes present in these metabolic pathways is varying.

ACKNOWLEDGEMENT

The authors are grateful to the SHIATS, Allahabad for providing the facilities and support to complete the present research work.

REFERENCES

Alves, R., Chaleil, R.A., Sternberg, M.J. 2002. Evolution of enzymes in metabolism: a network perspective. *J Mol Biol.,* **320:**751–770. doi: 10.1016/S0022-2836(02)00546-6.

Brocks, J.J., Logan, G.A., Buick, R., and Summons, R.E. 1999. Archean molecular fossils and the early rise of eukaryotes. *Science* 285: 1033-1036

Brown, J.R. and Doolittle, W.F. 1997. Archaea and the prokaryote-to-eukaryote transition. *Microbiol. Mol. Biol. Rev.* 61: 456-502

Lazcano A, Miller SL. 1996. The origin and early evolution of life: prebiotic chemistry, the pre-RNA world, and time. *Cell.* 1996;85:793–798. doi: 10.1016/S0092-8674 (00) 81263-5.

Lawrence JG, Roth JR. 1996. Selfish operons: horizontal transfer may drive the evolution of gene clusters. *Genetics,* **143:**1843–1860.

Singh, S., Singh, G., Sagar, N., Yadav, P. K., Jain P.A., Gautam, B. and Wadhwa, G. 2012. Insight into *Trichomonas vaginalis* genome evolution through metabolic pathways comparison. *Bioinformation* **8**(4):189-195

La Kanehisa M, Goto S, Kawashima S and Nakaya A. The *KEGG* databases at *GenomeNet. Nucleic Acids Res.* 2002;30:42–46. doi: 10.1093/nar/30.1.42.

8. Makarova, K.S., Aravind, L., Galperin, M.G., Grishin, N.V., Tatusov, R.L., Wolf, Y.I., and Koonin, E.V. 1999. Comparative genomics of the archaea (euryarchaeota): Evolution of conserved protein families, the stable core, and the variable shell. *Genome Res.* **9:** 608-628

Mojzsis, S.J., Arrhenius, G., Mckeegan, K.D., Harrison, T.M., Nutman, A.P., and Friend, C.R.L. 1996. Evidence for life on earth before 3,800 million

Tsoka S, Ouzounis CA. Functional versatility and molecular diversity of the metabolic map of *Escherichia coli. Genome Res.* 2001;11:1503–1510. doi: 10.1101/gr.187501.

Wang, Y.-C., Kumar, S., and Hedges, S.B. 1999. Divergence time estimates for the early history of animal phyla and the origin of plants, animals and fungi. *Proc. R. Soc. Lond. B* **266:**163-171