**Bioinformation**
www.bioinformation.net

*open access*
***Hypothesis***

# Genome subtraction for novel target definition in *Salmonella typhi*

**Bhawna Rathi*, Aditya N. Sarangi, Nidhi Trivedi**

Biomedical Informatics Centre, Sanjay Gandhi Postgraduate Institute of Medical Sciences, Lucknow 226014, India; Bhawna Rathi – Email: bhawna_rathi3@rediffmail.com; *Corresponding author

Received May 07, 2009; Revised August 24, 2009; Accepted September 11, 2009; Published October 11, 2009

**Abstract:**
Large genomic sequencing projects of pathogens as well as human genome leads to immense genomic and proteomic data which would be very beneficial for the novel target identification in pathogens. Subtractive genomic approach is one of the most useful strategies helpful in identification of potential targets. The approach works by subtracting the genes or proteins homologous to both host and the pathogen and identify those set of gene or proteins which are essential for the pathogen and are exclusively present in the pathogen. Subtractive genomic approach is employed to identify novel target in *salmonella typhi*. The pathogen has 4718 proteins out of which 300 are found to be essential ("indispensable to support cellular life") in the pathogen with no human homolog. Metabolic pathway analyses of these 300 essential proteins revealed that 149 proteins are exclusively involved in several metabolic pathway of *S. typhi*. 8 metabolic pathways are found to be present exclusively in the pathogen comprising of 27 enzymes unique to the pathogen. Thus, these 27 proteins may serve as prospective drug targets. Sub-cellular localization prediction of the 300 essential proteins was done which reveals that 11 proteins lie on the outer membrane of the pathogen which could be probable vaccine candidates.

**Keywords:** Subtractive genomics approach, proteome, drug targets.

**Background:**
The availability of large amount of genomic data generated by the microbial genomes and the human genome project has revolutionized the field of drug-discovery against threatening human pathogens **[1]**. These large sets of genomic data are useful in identification and characterization of the novel therapeutic targets and virulent factors prevalent in the pathogens. Subtractive genomic strategy is developed by assuming that the novel targets identified in the pathogen should be essential for the pathogen that is it should be involved in the replication, survival and a important component of various metabolic pathways and mechanisms occurring in the pathogen while at the same time should be absent on the host that is human and should have no homolog in human, so that when a drug or a lead compound is designed considering the potential target it should only be against the mechanism and functionality of the pathogen not the host. Subtractive genomics has been successfully used by authors to locate novel drug targets in *Pseudomonas aeruginosa* **[2]**. The work has been effectively complemented with the compilation of the Database of Essential Genes (DEG) for a number of pathogenic microorganisms **[3]**. The current studies make use of the subtractive genomics approach and DEG to analyze the complete genome of *Salmonella typhi* to search for potential vaccine candidates which would possibly lie on the surface membrane of the pathogen and drug targets.

*Salmonella enterica* serovar typhi is a human-specific gram-negative pathogen causing enteric typhoid fever, a severe infection of the reticuloendothelial system **[4], [5], [6]**. It has two strains CT18 (multiple drug resistant) **[7]** and Ty with a complete proteome of 4718 proteins. Worldwide, typhoid fever affects roughly millions of people annually, causing deaths. Infection of *S. typhi* leads to the development of typhoid, or enteric fever. This disease is characterized by the sudden onset of a sustained and systemic fever, severe headache, nausea, and loss of appetite. Other symptoms include constipation or diarrhea, enlargement of the spleen, possible development of meningitis, and/or general depression. Untreated typhoid fever cases result in mortality rates ranging from 12-30% while treated cases allow for 99% survival. The early administration of antibiotic treatment has proven to be highly effective in eliminating infections, but indiscriminate use of antibiotics has led to the emergence of multidrug-resistant strains of *S. enterica* serovar *Typhi* **[8]**. Chloramphenicol was the drug for the treatment of this infection till plasmid mediated chloramphenicol resistance was encountered [9]. Following this ciprofloxacin became the mainstay of treatment being a safer and more effective drug than Chloramphenicol but after clinical resistance to treatment with ciprofloxacin in the patients suffering from enteric fever, the choice left now is an expensive drug like ceftriaxone or cefexime.[10]. Resistance against ceftriaxone have been reported to CDC (Centre for Drug Control) [11] mild to moderate side effects have been shown for ceftriaxone. The novel targets identified by us using subtractive genomics will help enable understanding the biology of the pathogen to provide a more cost effective medication.

**Methodology:**
The systematic identification and characterization of potential targets in *salmonella typhi* is illustrated in **Figure 1**.

**Retrieval of proteomes of host and pathogen:**
The complete proteome of *Salmonella typhi* were retrieved from SwissProt **[12]** and protein sequences of *Homo sapiens* were downloaded from NCBI **[13]**. The Database of Essential genes was accessed from its location http://tubic.tju.edu.cn/deg/.

**Identification of essential proteins in *S. typhi:***
The *S. typhi* proteins were purged at 60% using CD-HIT **[14]** to identify the paralogs or duplicates proteins within the proteome of *S.typhi*. The paralogs are excluded and the remaining sets of protein were subjected to BlastP against *Homo sapiens* protein sequences with the expectation value (E-value) cutoff of $10^{-4}$. The resultant dataset obtained were with no homologs in *Homo sapiens*. BLASTP analysis was performed for the non homologous protein sequences of *S. typhi* against DEG with E-value cutoff score of $10^{-100}$. A minimum bit-score cut-off of 100 was used to screen out genes that appeared to represent essential genes. The protein sequences obtained are non homologous essential proteins of *S.typhi*.

**Metabolic pathway analysis:**
Metabolic pathway analysis of the essential proteins of S. typhi was done by KAAS server at KEGG for the identification of potential targets. **KAAS** (KEGG Automatic Annotation Server) provides functional annotation of genes by BLAST comparisons against the manually curated KEGG GENES database. The result contains KO (KEGG Orthology) assignments and automatically generated KEGG pathways. **[15]**

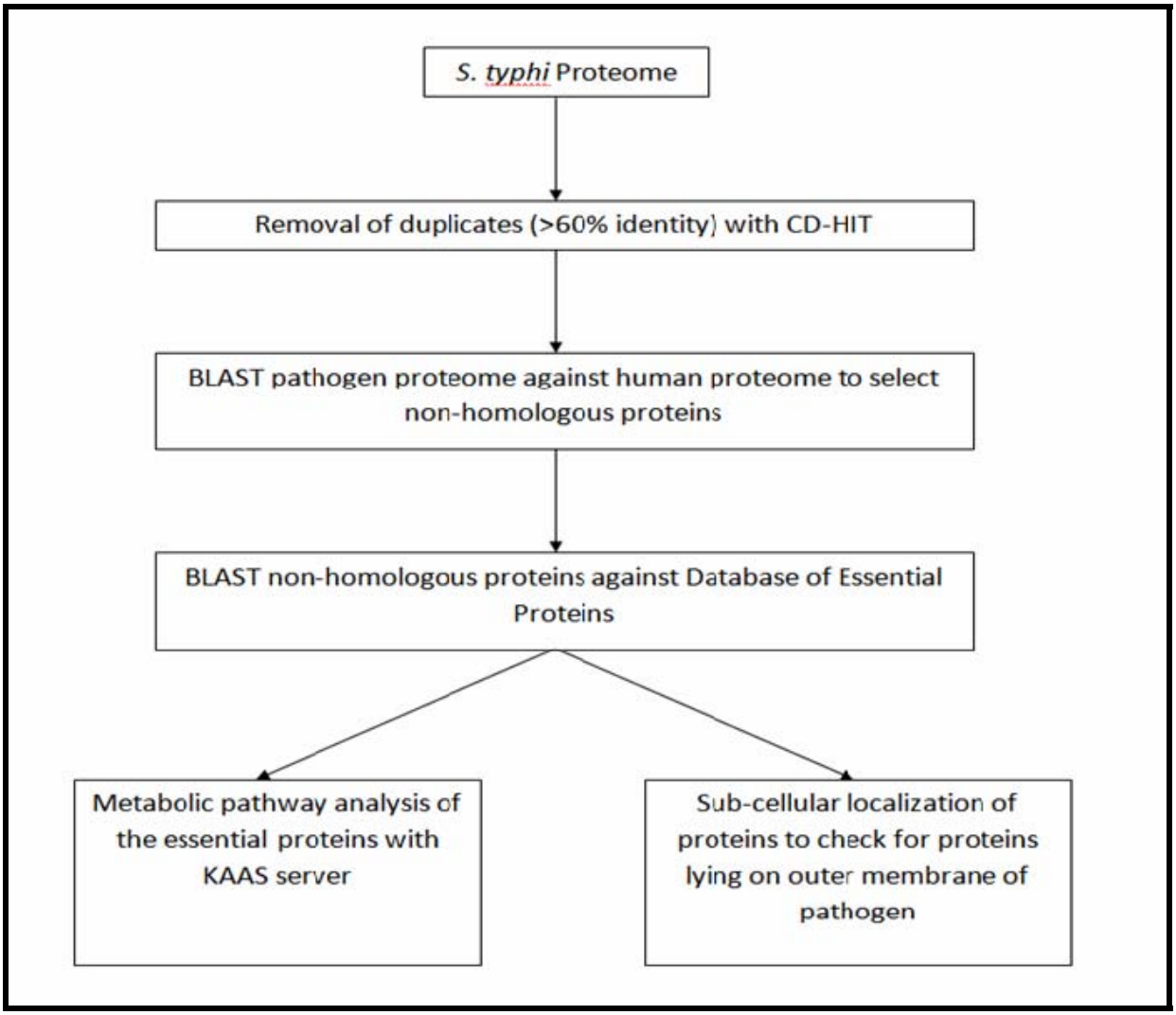


**Figure 1:** Flow chart for systematic identification and characterization of potential targets in *salmonella typhi.*

**Sub-cellular Localization prediction:**
Protein sub cellular localization prediction involves the computational prediction of where a protein resides in a cell. Prediction of protein sub cellular localization is an important component as it predicts the protein function and genome annotation, and it can aid the identification of targets. Sub-cellular localization analysis of the essential protein sequences has been done by Proteome Analyst Specialized Subcellular Localization Server v2.5 (PA-SUB) **[16]** to identify the surface membrane proteins which could be probable vaccine candidates.

**Discussion:**
The results obtained through computational analysis reveals that out of 4718 proteins in *salmonella typhi* 159 were identified as duplicates through CD-HIT with 60% similarity. The remaining 4559 paralogs were subjected to subtractive genomics which leads to 3570 proteins. These 3570 proteins when subjected to blastp against DEG database showed 300 proteins, which were essential for the pathogen. The results for subtractive proteome approach, metabolic pathway analysis and sub cellular localization are listed in Table No. 1(Supplementary material). The purpose of the present studies was to locate those essential proteins of *S. typhi* that play vital roles in the normal functioning of the bacterium within the host and to pick out them in the view of targeting. Detection of non-human homologs in the essential proteins of *S. typhi* with subsequent screening of the proteome to find the resultant protein product are likely to lead to development of drugs that exclusively interact with the pathogen. The non-human homologs of the surface proteins would represent potential vaccine candidates. 300 of the essential proteins were without human homologs. Metabolic pathway analyses of these 300 essential proteins by KAAS server at KEGG revealed that out of 300, 149 proteins might be concluded to be unique and are invariably linked with essential metabolic and signal transduction pathways. Presumably, screening against such novel targets for functional inhibitors will result in discovery of novel therapeutic compounds active against bacteria, including the increased number of antibiotic resistant clinical strains **[17]**.

Metabolic pathway analyses of the 149 essential proteins revealed that 15 proteins are involved in Carbohydrate Metabolism, 10 in Energy Metabolism, 5 in Lipid Metabolism, 4 in Nucleotide Metabolism, 30 in Amino Acid Metabolism, 20 in Glycan Biosynthesis and Metabolism, 16 in Metabolism of Co-factors and Vitamins, 20 in genetic information processing, 26 in environmental information processing and 2 in human disease. The results are summarized in **Table 2** (Supplementary material). Comparative

analysis of the metabolic pathways of the host (*Homo sapiens*) and the pathogen (*S.typhi*) by using Kyoto Encyclopedia of Genes and Genomes (KEGG) reveals 8 pathways which are unique to *S.typhi*. Thereafter, each selected pathway was screened for the unique enzymes and proteins involved. The peptidoglycan layer of the bacterial cell wall is the major structural element which plays an important role in pathogenesis as it provides resistance to osmotic lysis. D-alanine is the central molecule in the peptidoglycan assembly and cross-linking. D-alanine-D-alanine ligase (ddlA) is an important target as it is involved in D-alanine metabolism. Lipopolysaccharides (LPS) are also one of the main constituents of the outer cell wall of gram negative bacteria and play an important role for the survival of the pathogen. Out of the 14 enzymes involved in LPS biosyntheseis pathway, 13 enzymes are found to be essential for the variability of the bacteria and could be probable drug targets and it did not show homology with any human protein.

Two-component systems of bacteria represent the primary signal transduction paradigm in prokaryotic organisms. 8 essential enzymes were found to be potential targets in this pathway. Tryptophan synthase beta chain (trpB) is an important enzyme as it is involved in tyrosine and tryptophan biosynthesis pathway. Chemotaxis protein (MotA) and chemotaxis protein methyltransferase (CheR) is essential enzyme due to its involvement in multiple metabolic pathways like cell Motility, bacterial chemotaxis and flagellar assembly. Phosphoenolpyruvate (ppc) has been identified as a possible target due to its involvement in carbon fixation in photosynthetic organism, pyruvate metabolism and reductive carboxylase cycle. The focus of the present studies was to hunt for potential targets in *S. typhi* by computational approach. The sub-cellular localization prediction done by PA-SUB identify 11 proteins lying on the surface of the pathogen which could represent promising candidates for further characterization and analysis with a support to vaccine design. The results are summarized in Table No. 3 (Supplementary material)

**Conclusion:**

The availability of full genomic and proteomic sequences generated from the sequencing projects along with the computer-aided softwares to identify and characterize probable drug targets is a new emerging trend in pharmacogenomics . The application of the Database of essential genes helps to identify the potential drug targets in pathogens. The current study helps in the characterization of the potential proteins that could be targets for efficient drug design against *Salmonella typhi*. As subtractive genomic approach is applied for the identification of drug targets, so the drug would be specific for the pathogen and not lethal to the host. Molecular modeling of the targets will decipher the best possible active sites that can be targeted by simulations for drug design. Virtual screening against these potential targets might be useful in the discovery of potential therapeutic compounds against *Salmonella typhi*.

**References:**

**[1]** L Miesel *et al., Nat Rev Genet.* 4: 442 (2003) [PMID: 12776214]
**[2]** K Sakharkar *et al., In Silico Biol.* 4: 28 (2004) [PMID: 15724285]
**[3]** R Zhang *et al., Nucleic Acids Res.* 37: D455 (2009) [PMID: 18974178]
**[4]** P Everest e*t al.,Trends Microbiol.* 9: 316 (2001) [PMID: 11435104]
**[5]** JE Galan *Mol Microbiol.* 20: 263 (1996) [PMID: 8733226]
**[6]** BD Jones *et al., Annu Rev Immunol.*14: 533 (1996) [PMID: 8717524]
**[7]** J Parkhill *et al., Nature* 413: (2001) [PMID: 11677608]
**[8]** B Rowe *et al., Clin Infect Dis.* 24**:** S106 (1997) [PMID: 8994789]
**[9]** A Kapil *et al., Indian J Pathol Microbiol.* 37: 179 (1994) [PMID: 7959985]
**[10]** A Kapil *Indian J Med Res.* 121: 83 (2005) [PMID: 15756040]
**[11]** E. Steinburg *et al.* Antimicrobial Resistance of *Salmonella typhi* in the United States: the National Antimicrobial Monitoring System (NARMS), 1999
**[12]** http://www.expasy.ch/sprot/
**[13]** http://www.ncbi.nlm.nih.gov/
**[14]** W Li *et al., Bioinformatics* 17: 282 (2001) [PMID: 11294794]
**[15]** Y Moriya *et al., Nucleic Acids Res.* 35: W182 (2007) [PMID: 17526522]
**[16]** Z Lu *et al.,* Bioinformatics 20: 547 (2004) [PMID: 1499045]
**[17]** J Thanassi *et al., Nucleic Acid Res.* 30**:** 3152 (2002) [PMID: 12136097]

**Edited by P. Kangueane**
**Citation: Rathi *et al.*,** Bioinformation 4(4): 143-150 (2009)

**Bioinformation**
www.bioinformation.net

*open access*
***Hypothesis***

# Supplementary material

**Table 1:** Substractive proteomic and metabolic pathway analysis result for *Salmonella typhi*

| ***Salmonella typhi*** | **Number** |
|---|---|
| Total Number of proteins | 4718 |
| Duplicates (>60% identical) in CD-HIT | 159 |
| Non-paralogs | 4559 |
| Non-human homologous proteins (E-value $10^{-4}$) | 3570 |
| Essential protein in DEG (E-value $10^{-100}$) | 300 |
| Essential proteins involved in metabolic pathways | 149 |
| Pathways unique to the organism (*S.typhi*) | 8 |
| Proteins involved in unique pathways | 27 |
| Membrane associated non-human homologs of essential genes | 11 |

**Table 2:** Essential proteins of *S.typhi* involved in several metabolic pathways

| SN | KO | Protein Name | Gene Name | Pathway | EC |
|---|---|---|---|---|---|
| **Metabolism** | | | | | |
| **Carbohydrate metabolism** | | | | | |
| 1 | K02777 | glucose-specific IIA component | crr | Phosphotransfe-rase system | EC:2.7.1.69 |
| 2 | K01643 | citrate lyase subunit alpha | citF | Environmental Information Processing | EC:4.1.3.6 |
| 3 | K00117 | Quinoprotein dehydrogenase glucose | gcd | Pentose pathway phosphate | EC:1.1.1.130 |
| 4 | K08092 | 3-dehydro-L-gulonate 2-Dehydrogenase | E1.1.1.130 | Pentose and glucuronate Interconversions | EC:1.1.1.130 |
| 5 | K02798 | mannitol-specific IIA component | mtlA | Phosphotransfe-rase system | EC:2.7.1.69 |
| 6 | K01818 | L-fucose isomerase | fucI | Fructose and mannose metabolism | EC 5.3.1.25 |
| 7 | K02821 | ascorbate-specific IIA component | sgaA | Phosphotransfe-rase system | EC:2.7.1.69 |
| 8 | K01788 | -acylglucosamine-6-phosphate 2-epimerase | nanE | Aminosugars metabolism | EC:5.1.3.9 |
| 9 | K03431 | phosphoglucosamine mutase | glmM | Aminosugars metabolism | EC:5.4.2.10 |
| 10 | K00790 | UDP-N-acetylglucosamine 1-carboxyvinyltransferase | murA | Aminosugars metabolism | EC:2.5.1.7 |
| 11 | K00075 | UDP-N-acetylmuramate dehydrogenase | murB | Aminosugars metabolism | EC:1.1.1.158 |
| 12 | K01595 | phosphoenolpyruvate carboxylase | Ppc | Pyruvate metabolism | EC:4.1.1.31 |
| 13 | K00656 | formate C-acetyltransferase | pflD | Pyruvate metabolism | EC:2.3.1.54 |
| 14 | K00925 | acetate kinase | ackA | Pyruvate metabolism | EC:2.7.2.1 |
| 15 | K00932 | propionate kinase | tdcD | Propanoate metabolism | EC:2.7.2.15 |
| **Energy metabolism** | | | | | |
| 1 | K00425 | cytochrome bd-I oxidase subunit I | cydA | Oxidative phosphorylation | EC:1.10.3 |
| 2 | K00426 | cytochrome bd-I oxidase subunit I | cydB | Oxidative phosphorylation | EC:1.10.3 |
| 3 | K01595 | phosphoenolpyruvate carboxylase | Ppc | Pyruvate metabolism | EC:4.1.1.31 |
| 4 | K00926 | carbamate kinase | arc | Nitrogen metabolism | EC:2.7.2.2 |
| 5 | K01916 | NAD+ synthase | NadE | Nitrogen metabolism | EC:6.3.1.5 |
| 6 | K01914 | aspartate--ammonia ligase | AsnA | Nitrogen metabolism | EC:6.3.1.1 |
| 7 | K00264 | Glutamate synthase (NADPH/NADH) | GLT1 | Nitrogen metabolism | EC:1.4.1.13 1.4.1.14 |
| 8 | K03385 | formate-dependent nitrite reductase | NrfA | Nitrogen metabolism | EC:1.7.2.2 |
| 9 | K00369 | nitrate reductase | E1.7.99.4 | Nitrogen metabolism | EC:1.7.99.4 |
| 10 | K00640 | serine O-acetyltransferase | CysE | Sulfur metabolism | EC:2.3.1.30 |
| **Lipid metabolism** | | | | | |
| 1 | K00648 | 3-oxoacyl-[acyl-carrier-protein] synthase III | fabH | Fatty acid biosynthesis | EC:2.3.1.180 |
| 2 | K03527 | 4-hydroxy-3-methylbut-2-enyl diphosphate reductase | ispH | Biosynthesis of steroids | EC:1.17.1.2 |
| 3 | K03526 | (E)-4-hydroxy-3-methylbut- 2-enyl-diphosphate synthase | ispG | Biosynthesis of steroids | EC:1.17.7.1 |
| 4 | K00919 | 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase | ispE | Biosynthesis of steroids | EC:2.7.1.148 |
| 5 | K00099 | 1-deoxy-D-xylulose-5-phosphate reductoisomerase | Dxr | Biosynthesis of steroids | EC:1.1.1.267 |

| | | | | | |
|---|---|---|---|---|---|
| **Nucleotide metabolism** | | | | | |
| 1 | K00951 | GTP pyrophosphokinase | relA | Purine metabolism | EC:2.7.6.5 |
| 2 | K09903 | uridylate kinase | pyrH | Pyrimidine metabolism | EC:2.7.4.22 |
| 3 | K03040 | DNA-directed RNA polymerase subunit alpha | rpoA | Genetic Information Processing | EC:2.7.7.6 |
| 4 | K02319 | DNA polymerase I | polB1 | Purine metabolism | EC:2.7.7.7 |
| **Amino acid metabolism** | | | | | |
| 1 | K00926 | carbamate kinase | arc | Glutamate metabolism | EC:2.7.2.2 |
| 2 | K01776 | glutamate racemase | murI | Glutamate metabolism | EC:5.1.1.3 |
| 3 | K01775 | alanine racemase | Alr | Alanine and aspartate metabolism | EC:5.1.1.1 |
| 4 | K01270 | aminoacylhistidine dipeptidase | pepD | Alanine and aspartate metabolism | EC:3.4.13.3 |
| 5 | K00003 | homoserine dehydrogenase | thrA | Glycine,serine and threonine metabolism | EC:1.1.1.3 |
| 6 | K00133 | aspartate-semialdehyde dehydrogenase | asd | Glycine,serine and threonine metabolism | EC:1.2.1.11 |
| 7 | K00549 | 5- methyltetrahydropteroyltriglutamate—homocysteine | metE | Methionine metabolism | EC:2.1.1.14 |
| 8 | K01243 | S-adenosylhomocysteine/5'-methylthioadenosine nucleosidase | mtnN, mtn,pfs | Methionine metabolism | EC:3.2.2.9 |
| 9 | K00215 | dihydrodipicolinate reductase | dapB | Lysine biosynthesis | EC:1.3.1.26 |
| 10 | K00674 | 2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase | dapD | Lysine biosynthesis | EC:2.3.1.117 |
| 11 | K01778 | diaminopimelate epimerase | dapF | Lysine biosynthesis | EC:5.1.1.7 |
| 12 | K01929 | UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate--D-alanyl- D-alanine ligase | murF | Lysine biosynthesis | EC:6.3.2.10 |
| 13 | K01928 | UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase | murE | Lysine biosynthesis | EC:6.3.2.13 |
| 14 | K01484 | succinylarginine dihydrolase | astB | Arginine and proline metabolism | EC:3.5.3.23 |
| 15 | K00673 | arginine N-succinyltransferase | astA | Arginine and proline metabolism | EC:2.3.1.109 |
| 16 | K00765 | ATP phosphoribosyltransferase | hisG | Histidine metabolism | EC:2.4.2.17 |
| 17 | K01523 | phosphoribosyl-ATP pyrophosphohydrolase | hisE | Histidine metabolism | EC:3.6.1.31 |
| 18 | K01496 | phosphoribosyl-AMP cyclohydrolase | hisI | Histidine metabolism | EC:3.5.4.19 |
| 19 | K01693 | imidazoleglycerol-phosphate dehydratase | hisB | Histidine metabolism | EC:4.2.1.19 |
| 20 | K01089 | imidazoleglycerol-phosphate dehydratase / histidinol-hosphatase | hisB | Histidine metabolism | EC:4.2.1.19 3.1.3.15 |
| 21 | K01626 | 3-deoxy-7-phosphoheptulonate synthase | aroF,aroG, aroH | Phenylalanine, tyrosine and tryptophan biosynthesis | EC:2.5.1.54 |
| 22 | K01735 | 3-dehydroquinate synthase | ARO1 | Phenylalanine, tyrosine and tryptophan biosynthesis | EC:4.2.3.4 |
| 23 | K01696 | Tryptophan synthase beta chain | trpB | Phenylalanine, tyrosine and tryptophan biosynthesis | EC:4.2.1.20 |
| 24 | K01695 | Tryptophan synthase alpha chain | trpA | Phenylalanine, tyrosine and tryptophan biosynthesis | EC:4.2.1.20 |
| 25 | K01736 | chorismate synthase | aroC | Phenylalanine, tyrosine and tryptophan biosynthesis | EC:4.2.3.5 |
| 26 | K01850 | chorismate mutase | E5.4.99.5 | Phenylalanine, tyrosine and tryptophan biosynthesis | EC:5.4.99.5 |
| 27 | K00145 | N-acetyl-gamma-glutamyl- phosphate reductase | argC | Urea cycle And metabolism of amino groups | EC:1.2.1.38 |
| **Metabolism of another amino acids** | | | | | |
| 1 | K01925 | UDP-N-acetylmuramoylalanine--D-glutamate ligase | murD | D-Glutamine and D-glutamate metabolism | EC:6.3.2.9 |
| 2 | K01924 | UDP-N-acetylmuramate--alanine ligase | murC | D-Glutamine and D-glutamate metabolism | EC:6.3.2.8 |
| 3 | K01921 | D-alanine-D-alanine ligase | ddlA | D-Alanine metabolism | EC:6.3.2.4 |
| **Glycan biosynthesis and metabolism** | | | | | |
| 1 | K00677 | UDP-N-acetylglucosamine | lpxA | Lipopolysaccharide | EC:2.3.1.129 |

| | | | | | |
|---|---|---|---|---|---|
| | | acyltransferase | | biosynthesis | |
| 2 | K02535 | UDP-3-O-[3-hdroxymyristoyl] N- acetylglucosamine deacetylase | lpxC | Lipopolysaccharide biosynthesis | EC:3.5.1. |
| 3 | K02536 | UDP-3-O-[3- hydroxymyristoyl] glucosamine N-acyltransferase | lpxD | Lipopolysaccharide biosynthesis | EC:2.3.1. |
| 4 | K03269 | UDP-2,3-diacylglucosamine hydrolase | lpxH | Lipopolysaccharide biosynthesis | EC:3.6.1. |
| 5 | K00748 | lipid-A-disaccharide synthase | lpxB | Lipopolysaccharide biosynthesis | EC:2.4.1.182 |
| 6 | K00912 | tetraacyldisaccharide 4'-kinase | lpxK | Lipopolysaccharide biosynthesis | EC:2.7.1.130 |
| 7 | K02527 | 3-deoxy-D-manno-octulosonic-acid transferase | kdtA | Lipopolysaccharide biosynthesis | EC:2.-.-.- |
| 8 | K00979 | 3-deoxy-manno-octulosonate cytidylyltransferase | kdsB | Lipopolysaccharide biosynthesis | EC:2.7.7.38 |
| 9 | K01627 | 2-dehydro-3-deoxyphosphooctonate aldolase | kdsA | Lipopolysaccharide biosynthesis | EC:2.5.1.55 |
| 10 | K02841 | heptosyltransferase I | waaC, rfaC | Lipopolysaccharide biosynthesis | EC:2.4.-.- |
| 11 | K02843 | heptosyltransferase II | waaF, rfaF | Lipopolysaccharide biosynthesis | EC:2.4.-.- |
| 12 | K02840 | Galactosyltransferase | waaB, rfaB | Lipopolysaccharide biosynthesis | EC:2.4.1.- |
| 13 | K02844 | Glucosyltransferase | waaG, rfaG | Lipopolysaccharide biosynthesis | EC:2.4.1.- |
| 14 | K02847 | O-antigen ligase | waaL, rfaL | Lipopolysaccharide biosynthesis | EC:6.-.-.- |
| 15 | K01921 | D-alanine-D-alanine ligase | ddlA | Peptidoglycan biosynthesis | EC:6.3.2.4 |
| 16 | K01000 | phospho-N-acetylmuramoyl-pentapeptide-transferase | mraY | Peptidoglycan biosynthesis | EC:2.7.8.13 |
| 17 | K02563 | UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase | murG | Peptidoglycan biosynthesis | EC:2.4.1.227 |
| 18 | K01924 | UDP-N-acetylmuramate-alanine ligase | murC | Peptidoglycan biosynthesis | EC:6.3.2.8 |
| 19 | K01925 | UDP-N-acetylmuramoylalanine--D-glutamate ligase | murD | Peptidoglycan biosynthesis | EC:6.3.2.9 |
| 20 | K03587 | cell division protein FtsI | ftsI | Peptidoglycan biosynthesis | EC:2.4.1.129 |
| **Metabolism of Co-factors and Vitamins** | | | | | |
| 1 | K03147 | thiamine biosynthesis protein ThiC | thiC | Thiamine metabolism | |
| 2 | K00946 | thiamine-monophosphate kinase | thiL | Thiamine metabolism | EC:2.7.4.16 |
| 3 | K01497 | GTP cyclohydrolase II | ribA | Riboflavin metabolism | EC:3.5.4.25 |
| 4 | K01498 | diaminohydroxyphosphoribosylaminopyrimidine deaminase | E3.5.4.26 | Riboflavin metabolism | EC:3.5.4.26 |
| 5 | K00082 | 5-amino-6-(5-phosphoribosylamino) uracil reductase | E1.1.1.193 | Riboflavin metabolism | EC:1.1.1.193 |
| 6 | K02858 | 3,4-dihydroxy 2-butanone 4-phosphate synthase | ribB | Riboflavin metabolism | |
| 7 | K00793 | riboflavin synthase alpha chain | ribE | Riboflavin metabolism | EC:2.5.1.9 |
| 8 | K03474 | pyridoxine synthase 5-phosphate | pdxJ | Riboflavin metabolism | EC:2.6.99.2 |
| 9 | K00969 | nicotinate-nucleotide adenylyltransferase | nadD | Nicotinate and nicotinamide metabolism | EC:2.7.7.18 |
| 10 | K03517 | quinolinate synthase | nadA | Nicotinate and nicotinamide metabolism | |
| 11 | K01012 | biotin synthetase | bioB | Biotin metabolism | EC:2.8.1.6 |
| 12 | K01664 | para-aminobenzoate synthetase component II | pabA | Folate biosynthesis | EC:2.6.1.85 |
| 13 | K02302 | uroporphyrin-III C- methyltransferase / precorrin-2 dehydrogenase / sirohydrochlorin ferrochelatase | cysG | Porphyrin and chlorophyll metabolism | EC:2.1.1.107 1.3.1.76 4.99.1.4 |
| 14 | K02492 | glutamyl-tRNA reductase | hemA | Porphyrin and chlorophyll metabolism | EC:1.2.1.70 |
| 15 | K02551 | 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase | mend | Ubiquinone and menaquinone biosynthesis | EC:2.2.1.9 |
| 16 | K03182 | 3-octaprenyl-4-hydroxybenzoate carboxy- lyase UbiD | ubiD | Ubiquinone and menaquinone biosynthesis | EC:4.1.1.- |
| **Xenobiotics Biodegradation and Metabolism** | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| 1 | K06281 | hydrogenase large subunit | E1.12.99.6L | Xenobiotics and biodegradation metabolism | EC:1.12.99.6 |
| **Genetic information processing** | | | | | |
| **Transcription** | | | | | |
| 1 | K03040 | DNA-directed RNA polymerase subunit alpha | rpoA | RNA polymerase | EC:2.7.7.6 |
| **Translation** | | | | | |
| 1 | K02986 | small subunit ribosomal RP-S4, rotein S4 | rpsD | Translation | |
| 2 | K01878 | glycyl-tRNA synthetase alpha chain | glyQ | Aminoacyl-tRNA biosynthesis | EC:6.1.1.14 |
| **Folding, sorting and degradation** | | | | | |
| 1 | K03070 | Preprotein translocase SecA subunit | secA | Folding, Sorting and Degradation | |
| 2 | K03076 | preprotein translocase SecY subunit | secY | Folding, Sorting and Degradation | |
| 3 | K03072 | Preprotein translocase SecD subunit | secD | Folding, Sorting and Degradation | |
| 4 | K03074 | preprotein translocase SecF subunit | secF | Folding, Sorting and Degradation | |
| **Replication and Repair** | | | | | |
| 1 | K02342 | DNA polymerase III subunit DPO3E, epsilon | dnaQ | DNA replication | EC:2.7.7.7 |
| 2 | K02337 | DNA polymerase III subunit DPO3A1, alpha | dnaE | DNA Replication | EC:2.7.7.7 |
| 3 | K02341 | DNA polymerase III subunit DPO3D2, delta | holB | DNA Replication | EC:2.7.7.7 |
| 4 | K02338 | DNA polymerase III subunit DPO3B, beta | dnaN | DNA Replication | EC:2.7.7.7 |
| 5 | K02340 | DNA polymerase III subunit DPO3D1, delta | holA | DNA Replication | EC:2.7.7.7 |
| 6 | K02314 | replicative DNA helicase | dnaB | DNA Replication | EC:3.6.1.- |
| 7 | K02316 | DNA primase | dnaG | DNA Replication | EC:2.7.7.- |
| 8 | K03657 | DNA helicase II / ATP-dependent DNA helicase PcrA | uvrD, pcrA | Nucleotide excision repair | EC:3.6.1.- |
| 9 | K01141 | exodeoxyribonuclease I | sbcB | Mismatch repair | EC:3.1.11.1 |
| 10 | K03582 | exodeoxyribonuclease V beta subunit | recB | Homologous recombination | EC:3.1.11.5 |
| 11 | K03583 | exodeoxyribonuclease gamma subunit V | recC | Homologous recombination | EC:3.1.11.5 |
| 12 | K03629 | DNA replication and repair protein RecF | recF | Homologous recombination | EC:3.1.11.5 |
| 13 | K04066 | primosomal protein N' | priA | Homologous recombination | EC:3.6.1.- |
| **Environmental Information Processing** | | | | | |
| **Membrane Transport** | | | | | |
| 1 | K02047 | sulfate transport system permease protein | cysW | ABC transporters | |
| 2 | K11070 | spermidine/putrescine transport system permease protein | potC | ABC Transporters | |
| 3 | K11069 | spermidine/putrescine transport system substrate-binding protein | potD | ABC transporters | |
| 4 | K10540 | methyl-galactoside transport system protein | mglB | ABC Transporters | |
| 5 | K02040 | phosphate transport system substrate-binding protein | pstS | ABC Transporters | |
| 6 | K10015 | histidine transport system permease protein | hisM | ABC Transporters | |
| 7 | K10002 | glutamate/aspartate transport system permease protein | gltK | ABC Transporters | |
| 8 | K10009 | cystine transport system permease protein | ABC.CYST.P | ABC Transporters | |
| 9 | K02035 | peptide/nickel transport system substrate-binding protein | ABC.PE.S | ABC Transporters | |
| 10 | K02016 | iron complex transport system substrate-binding protein | ABC.FEV.S | ABC Transporters | |
| 11 | K09808 | lipoprotein-releasing system permease protein | ABC.LPT.P, lolC, lolE | ABC Transporters | |
| 12 | K09811 | cell division transport system permease protein | ftsX | ABC Transporters | |
| 13 | K07091 | lipopolysaccharide export system permease protein | lptF | ABC Transporters | |
| 14 | K11720 | lipopolysaccharide export system permease protein | lptG | ABC transporters | |
| 15 | K02778 | PTS system, glucose-specific IIB component | PTS-Glc-EIIB, ptsG | | |
| 16 | K03475 | PTS system, ascorbate-specific IIC component | PTS-Ula-EIIC, laA, sgaT | | |

| | | | | | |
|---|---|---|---|---|---|
| 17 | K08484 | phosphotransferase system, enzyme I, PtsP | PTS-EI.PTSP, ptsP | | |
| **Signal Transduction** | | | | | |
| 1 | K07636 | two-component system, OmpR family, phosphate regulon sensor histidine kinase PhoR | PhoR | Signal Transduction | EC:2.7.13.3 |
| 2 | K07639 | two-component system, OmpR family, sensor histidine kinase RstB | RstB | Signal Transduction | EC:2.7.13.3 |
| 3 | K02556 | chemotaxis protein MotA | motA | Signal Transduction | |
| 4 | K00370 | nitrate reductase 1, alpha subunit | narG | Signal Transduction | EC:1.7.99.4 |
| 5 | K00990 | [protein-PII] uridylyltransferase | glnD | Signal Transduction | EC:2.7.7.59 |
| 6 | K03407 | two-component system, chemotaxis family, sensor kinase CheA | cheA | Signal Transduction | EC:2.7.13.3 |
| 7 | K00575 | chemotaxis protein methyltransferase CheR | cheR | Signal Transduction | EC:2.1.1.80 |
| **Human Diseases** | | | | | |
| **Infectious Diseases** | | | | | |
| 1 | K03092 | RNA polymerase sigma-54 factor | SIG54, rpoN | Vibrio cholerae phathogenic cycle | |
| 2 | K05851 | adenylate cyclase, class 1 | E4.6.1.1A, cyaA | Vibrio cholerae phathogenic cycle | EC:4.6.1.1 |

**Table 3:** List of the outer membrane proteins of *Salmonlla typhi* identified by PA-SUB

| S.N | Accession No | Name of Protein | Sub-Cellular Localization |
|---|---|---|---|
| 1 | Q56110 | Outer membrane prorein S1 | Outer membrane |
| 2 | Q56119 | Outer membrane pore protein | Outer membrane |
| 3 | Q8Z8P3 | Outer membrane usher protein FimD | Outer membrane |
| 4 | Q8Z944 | Outer membrane fimbrial usher protein | Outer membrane |
| 5 | Q8Z4Y8 | Long chain fatty acid transport protein | Outer membrane |
| 6 | Q8Z1S4 | Putative Type-I section protein | Outer membrane |
| 7 | Q8XEL5 | Putative exported protein | Outer membrane |
| 8 | Q8Z9A3 | Outer membrane protein assembly factor yaeT | Outer membrane |
| 9 | Q8Z9J6 | LPS-assembly protein | Outer membrane |
| 10 | Q8Z4J0 | Putative lipoprotein | Outer membrane |
| 11 | Q8Z6A0 | Outer membrane lipoprotein lolB | Outer membrane |