# Initial automated reconstruction

## Pathway Tools

PathoLogic component of Pathway Tools infers metabolic pathways by analyzing the genome annotation with respect to a reference database of metabolic pathways, MetaCyc [1](#_ENREF_1). It also generates reports that summarize the evidence used for deducing the pathways inferred and the results of comparing the reactions carried out in the subject organism versus those in the reference database. The major steps to create draft reconstruction using Pathologic component of Pathway Tool version 17.0 are:

*Preparation of input data files*: The genome sequence and annotation files for *L. donovani* BPK282A1 were retrieved in FASTA and genebank format respectively to generate input file. Replicon editor was used to submit the sequence and annotation file for all 36 chromosome of *L. donovani* genome. The genome-scale metabolic network of *L. major* was provided as reference network since *L. major* and *L. donovani* are phylogenetic neighbors and genome sequences of both species are very similar.

*Building draft reconstruction*: At first, to test whether the PathoLogic is able to parse the input file properly or there is error, trial parse operation was performed. Parsing errors were removed and finally draft reconstruction operation was performed. This initial build provides useful information and was used as a basis for further manual curation process. By analyzing draft reconstruction it can be concluded that the reconstructed network is not complete and has many metabolic reaction gaps.

**Figure S1. Cellular overview diagram for L. donovani BPK282A1 draft reconstruction.** The icon shapes denotes compound classes i.e. Square: carbohydrate, Triangle: amino acid, upside down triangle: cofactor, ellipse (horizontal): purine, ellipse (vertical): pyrimidine

## RAVEN Toolbox

First of all KEGG reaction database was downloaded and a metabolic model of all reactions was generated. Then Protein sequences for each KO are retrieved and aligned using MUSCLE. In the next step Hidden Markov model (HMMs) is generated based on the sequences for each KO using HMMER. Now set of HMMs are queried with the protein sequences of the organism of interest. Finally, GEM for the *L. donovani* BPK282A1 is constructed by linking the reactions with the corresponding genes.

The first step to create a draft network was to import the *L. major* model in RAVEN toolbox using '*importModel'*. However, standardization of model was performed prior to its import as the model had inconsistencies in metabolite naming and model structure. After successful parsing of model, a Bi-directional Blastp was carried out between *L.major* and *L.donovani* protein fasta files using '*getBlast*'. The resulting blast Structure was further used by '*getModelFromHomology*' to create a draft metabolic network of *L. donovani.* The draft network was unable to produce biomass as it lacks exchange and transport reaction and has many metabolic gaps.

**Table S1. Characteristics of draft networks**

|  |  |  |  |
| --- | --- | --- | --- |
| **Components** | **Pathway Tools** | **GEMSiRV-MrBac** | **RAVEN Toolbox (KEGG)** |
| Total genes | 8032 | 8032 | 8032 |
| Metabolic genes | 809 | 485 | 398 |
| Pathways | 109 | 58 | 61 |
| Enzymatic reactions | 977 | 659 | 680 |
| Transport reactions | 05 | 28 | 0 |
| Metabolites | 738 | 938 | 870 |
| Exchange reactions | 0 | 32 | 0 |
| Compartments | 2 | 8 | 1 |

# Experimental characterization of biomass constituents

## Biomass Lyophilisation

*L. donovani* wild type (WT, MHOM/80/IN/Dd8) promastigotes were cultured at 24 °C in RPMI 1640 medium (Sigma). 20 ml culture was taken and centrifuged at 6000x g, 4 °C10 min and the pellets were washed twice in PBS, pH 7.4 to get a pellet free of culture medium. The recovered pellet was kept at -80 °C deep freezer followed by lyophilization under vacuum using Lyophilizer (Allied Frost).

## Biomass total protein content

Total protein content was determined by Coomassie blue (Bio-Rad protein assay) as described in [2](#_ENREF_2) from a solution of 1.5 mg lyophilized biomass. 1 ml biomass sample was incubated at 40 °C in 0.1 M NaOH/ 0.1 % SDS. Assay was performed by taking supernatant after centrifugation at high speed. Freshly prepared BSA in 0.1 M NaOH/ 0.1 % SDS was taken as standard.

## Biomass RNA content

RNA content of biomass was determined by conventional TRIzol-based RNA isolation method [3](#_ENREF_3) using TRI reagent (Sigma) according to the manufacturer’s protocol. Concentration of RNA was determined by using NanoPhotometer (Implen).

## Biomass DNA content

DNA content of biomass was determined by genomic DNA isolation kit (Promega) according to manufacturer’s protocol. Concentration of DNA was determined by using NanoPhotometer (Implen).

# Estimation of biomass equation

After determination of protein, DNA and RNA content, calculation for coefficient of was performed.

**Table S2. Molar estimation of each amino acid per gram DW of cell**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Cellular content of Protein** | **Amino acids** | **Fractional contribution** | **Molecular weight monomer** | **gram monomer/mol protein** | **gram protein/mol protein** | **gram monomer/gram protein** | **mol monomer/gDW** | **mmol monomer/gDW** |
| **0.5090909** | Ala | 0.1207 | 71.0790 | 8.5794 | 108.5940 | 0.0790 | 0.0006 | 0.5659 |
|  | Arg | 0.0726 | 157.1970 | 11.4079 |  | 0.1051 | 0.0003 | 0.3402 |
|  | Asn | 0.0263 | 114.1040 | 2.9987 |  | 0.0276 | 0.0001 | 0.1232 |
|  | Asp | 0.0487 | 114.0800 | 5.5557 |  | 0.0512 | 0.0002 | 0.2283 |
|  | Cys | 0.0187 | 103.1450 | 1.9250 |  | 0.0177 | 0.0001 | 0.0875 |
|  | Gln | 0.0410 | 128.1310 | 5.2550 |  | 0.0484 | 0.0002 | 0.1923 |
|  | Glu | 0.0594 | 128.1070 | 7.6062 |  | 0.0700 | 0.0003 | 0.2783 |
|  | Gly | 0.0652 | 57.0520 | 3.7203 |  | 0.0343 | 0.0003 | 0.3057 |
|  | His | 0.0272 | 137.1420 | 3.7368 |  | 0.0344 | 0.0001 | 0.1277 |
|  | Ile | 0.0298 | 113.1600 | 3.3758 |  | 0.0311 | 0.0001 | 0.1399 |
|  | Leu | 0.0919 | 113.1600 | 10.4033 |  | 0.0958 | 0.0004 | 0.4310 |
|  | Lys | 0.0332 | 129.1830 | 4.2901 |  | 0.0395 | 0.0002 | 0.1557 |
|  | Met | 0.0226 | 131.1990 | 2.9668 |  | 0.0273 | 0.0001 | 0.1060 |
|  | Phe | 0.0296 | 147.1770 | 4.3575 |  | 0.0401 | 0.0001 | 0.1388 |
|  | Pro | 0.0585 | 97.1170 | 5.6784 |  | 0.0523 | 0.0003 | 0.2741 |
|  | Ser | 0.0888 | 87.0780 | 7.7309 |  | 0.0712 | 0.0004 | 0.4162 |
|  | Thr | 0.0601 | 101.1050 | 6.0746 |  | 0.0559 | 0.0003 | 0.2817 |
|  | Trp | 0.0107 | 186.2140 | 1.9981 |  | 0.0184 | 0.0001 | 0.0503 |
|  | Tyr | 0.0240 | 163.1760 | 3.9122 |  | 0.0360 | 0.0001 | 0.1124 |
|  | Val | 0.0708 | 99.1330 | 7.0215 |  | 0.0647 | 0.0003 | 0.3321 |

Likewise, Table S2 and Table S3 show the calculation of molar composition for the nucleotides by using DNA and RNA composition *i.e.* 0.027 and 0.045 g/gDW respectively.

**Table S3. Molar estimation of each deoxy ribonucleotides per gram DW of cell**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Cellular content of DNA** | **nucleotide** | **Fractional contribution** | **Molecular weight monomer** | **gram monomer/mol DNA** | **gram DNA/mol DNA** | **gram monomer/gram DNA** | **mol monomer/gDW** | **mmol monomer/gDW** |
| 0.027 | dAMP | 20.1500 | 329.0700 | 6630.7605 | 32486.0515 | 0.2041 | 0.0000 | 0.0167 |
|  | dCMP | 29.8500 | 305.0600 | 9106.0410 |  | 0.2803 | 0.0000 | 0.0248 |
|  | dGMP | 29.8500 | 345.0600 | 10300.0410 |  | 0.3171 | 0.0000 | 0.0248 |
|  | dTMP | 20.1500 | 320.0600 | 6449.2090 |  | 0.1985 | 0.0000 | 0.0167 |

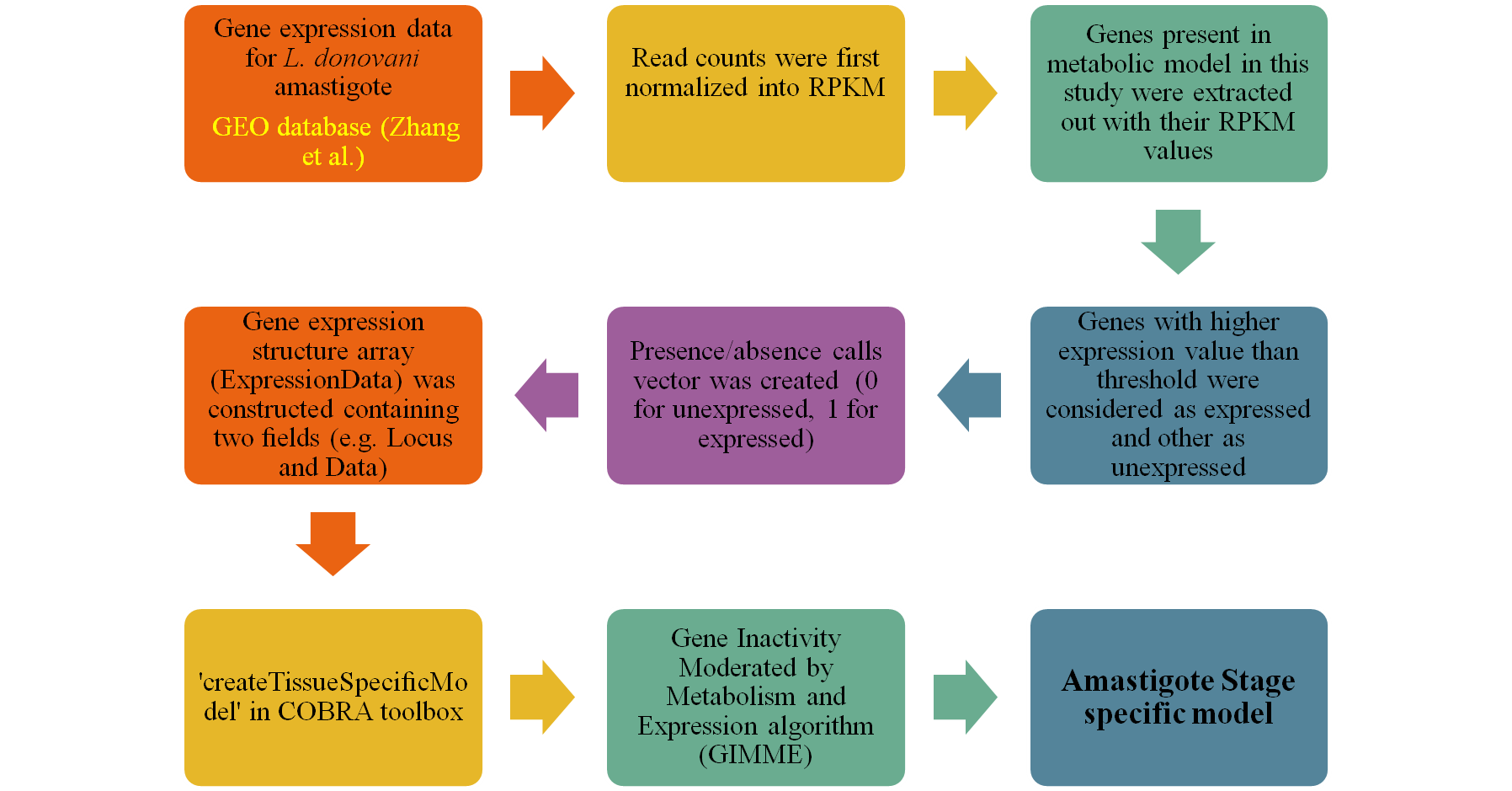
**Table S4. Molar estimation of each ribonucleotides per gram DW of cell**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Cellular content of RNA** | **nucleotide** | **Fractional contribution** | **Molecular weight monomer** | **gram monomer/mol RNA** | **gram DNA/mol RNA** | **gram monomer/gram RNA** | **mol monomer/gDW** | **mmol monomer/gDW** |
| 0.045 | AMP | 0.1937 | 345.0600 | 66.8523 | 337.3150 | 0.1982 | 0.0000 | 0.0258 |
|  | CMP | 0.2877 | 321.0500 | 92.3796 |  | 0.2739 | 0.0000 | 0.0384 |
|  | GMP | 0.2854 | 361.0600 | 103.0479 |  | 0.3055 | 0.0000 | 0.0381 |
|  | UMP | 0.2330 | 322.0400 | 75.0353 |  | 0.2224 | 0.0000 | 0.0311 |

These molar compositions of the macromolecules monomer were then added to the biomass equation. The modified biomass composition is as follows:

(0.5659) ala-L + (0.3402) arg-L + (0.1232) asn-L + (0.2283) asp-L + (0.0875) cys-L + (0.1923) gln-L + (0.2783) glu-L + (0.3057) gly + (0.1277) his-L + (0.1399) ile-L + (0.4310) leu-L + (0.1557) lys-L + (0.1060) met-L + (0.1388) phe-L + (0.2741) pro-L + (0.4162) ser-L + (0.2817) thr-L + (0.0503) trp-L + (0.1124) tyr-L + (0.3321) val-L + (0.0167) damp + (0.0248) dcmp + (0.0248) dgmp + (0.0167) dtmp + (0.0258) amp + (0.0384) cmp + (0.0381) gmp + (0.0311) ump + (0.0036) mannan + (0.0565) ergst + (0.0256) triglyc\_LM + (0.0123) zymst + (0.0013) 12dgr\_LM + (0.0015) mag\_LM + (0.0268) pe\_LM + (0.0620) pc\_LM + (0.0092) ptd1ino\_LM + (0.0021) clpn\_LM + (0.0333) ptrc + (0.0067) spmd + 0.000001 nad[c] + 0.000001 pnto\_R[c] + 0.000001 coa[c] + 0.000001 dolp[c] + (32.26) atp + (32.26) h2o --> (32.26) adp + (32.26) h + (32.26) pi

# Stage-specific metabolism



**Figure S2 Methodology for integration of gene expression data adopted in this study**

**Table S5 Amastigote stage specific reactions.**

|  |  |  |  |
| --- | --- | --- | --- |
| **Rxn name** | **Rxn description** | **Genes** | **Subsystem** |
| ASNN | L-asparaginase | LdBPK\_150440 | Alanine and Aspartate Metabolism |
| ASNNg | L-asparaginase (glycosome) | LdBPK\_364650 | Alanine and Aspartate Metabolism |
| BTMAT1m1 | Butyryl-[acyl-carrier protein]:malonyl-CoA C-acyltransferase | LdBPK\_340630 | Fatty Acid Biosynthesis |
| HEMAT2m2 | Hexanoyl-[acyl-carrier protein]:malonyl-CoA C-acyltransferase | LdBPK\_340630 | Fatty Acid Biosynthesis |
| OCMAT3m3 | Octanoyl-[acyl-carrier protein]:malonyl-CoA C-acyltransferase | LdBPK\_340630 | Fatty Acid Biosynthesis |
| DEMAT4m4 | Decanoyl-[acyl-carrier protein]:malonyl-CoA C-acyltransferase | LdBPK\_340630 | Fatty Acid Biosynthesis |
| DDMAT5m5 | Dodecanoyl-[acyl-carrier protein]: malonyl-CoA C-acyltransferase | LdBPK\_340630 | Fatty Acid Biosynthesis |
| MAT6m6 | Myristoyl-[acyl-carrier protein] synthase | LdBPK\_340630 | Fatty Acid Biosynthesis |
| ACACT7rm | Acetyl-CoA C-acyltransferase (tetradecanoyl-CoA) | LdBPK\_230860 | Fatty Acid Degradation |
| ECOAH7m | Trans-hexadec-2-enoyl-coA hydratase | LdBPK\_261530 LdBPK\_292420 LdBPK\_350360 LdBPK\_323830 LdBPK\_332740 LdBPK\_363440 | Fatty Acid Degradation |
| HACD7m | 3-hydroxyacyl-CoA dehydrogenase (3-oxohexadecanoyl-CoA) | LdBPK\_261530 LdBPK\_361200 LdBPK\_332740 | Fatty Acid Degradation |
| FAS160COAr | Fatty-acyl-CoA synthase (n-C16:0CoA) | LdBPK\_140710 LdBPK\_140740 LdBPK\_140750 LdBPK\_140760 | Fatty Acid Elongation |
| GTHS | Glutathione synthase | LdBPK\_140970 | Glutathione Metabolism |
| CDPDSP | CDPdiacylglycerol-serine O-phosphatidyltransferase (phosphatidylserine synthase) | LdBPK\_141280 | Glycerophospholipid metabolism |
| PAPAm | Phosphatidate phosphatase | LdBPK\_180440 LdBPK\_191380 | Glycerophospholipid metabolism |
| FBPg | Fructose-bisphosphatase, glycosomal | LdBPK\_041170 | Glycolysis/Gluconeogenesis |
| LDH\_D | D-lactate dehydrogenase | LdBPK\_271940 LdBPK\_290290 | Glyoxylase pathway |
| HCYSMT2 | Homocysteine S-methyltransferase | LdBPK\_366570 | Methionine Metabolism |
| NNAM | Nicotinamidase | LdBPK\_010470 | Nicotinate and Nicotinamide Metabolism |
| HCO3Em3 | HCO3 equilibration reaction |  | Other |
| XYLTRED\_D | D-xylulose reductase | LdBPK\_330530 | Pentose and Glucuronate Interconversions |
| G6PDH1 | Glucose 6-phosphate dehydrogenase | LdBPK\_340080 | Pentose Phosphate Pathway |
| PGL | 6-phosphogluconolactonase | LdBPK\_262730 | Pentose Phosphate Pathway |
| GSS | Glutathionylspermidine synthase | LdBPK\_271770 | Polyamine metabolism |
| TRYS | Trypanothione synthetase | LdBPK\_230500 LdBPK\_364510 | Polyamine metabolism |
| ATPM | Nucleoside-triphosphatase (ATP maintenance requirment) | LdBPK\_364860 | Purine Metabolism |
| DCMPDA2 | dCMP deaminase |  | Pyrimidine Metabolism |
| DCYTD1 | Deoxycytidine deaminase | LdBPK\_170410 LdBPK\_322210 | Pyrimidine Metabolism |
| DURIK1m | Deoxyuridine kinase (ATP:Deoxyuridine) | LdBPK\_211450 | Pyrimidine Metabolism |
| ME1x | Malic enzyme (NAD) | LdBPK\_240790 | Pyruvate metabolism |
| C5STDSr | C-5 sterol desaturase | LdBPK\_310620 | Steroid Biosynthesis |
| C8STIr | C-8 sterol isomerase | LdBPK\_292250 | Steroid Biosynthesis |
| SAM24MTr | Sterol 24-c-methyltransferase | LdBPK\_362520 LdBPK\_362510 | Steroid Biosynthesis |

# References

1. R. Caspi, H. Foerster, C. A. Fulcher, P. Kaipa, M. Krummenacker, M. Latendresse, S. Paley, S. Y. Rhee, A. G. Shearer and C. Tissier, *Nucleic Acids Res.*, 2008, **36**, D623-D631.

2. A. P. Andersen and P. Hellung‐Larsen, *J. Cell. Biochem.*, 1989, **41**, 125-133.

3. P. Chomczynski and N. Sacchi, *Anal. Biochem.*, 1987, **162**, 156-159.