**Supplementary Table 1.** Pathway comparison data of *C. parvum* (*Cpv*) and human collected from KEGG pathway database

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sl. No. | Pathway ID | Name of *Cpv* pathways | Found in Human | Genes Associated (Mentioned for only those pathways not found in human) |
| 1 | cpv00010 | Glycolysis / Gluconeogenesis | Yes |  |
| 2 | cpv00030 | Pentose phosphate pathway | Yes |  |
| 3 | cpv00040 | Pentose and glucuronate interconversions | Yes |  |
| 4 | cpv00051 | Fructose and mannose metabolism | Yes |  |
| 5 | cpv00052 | Galactose metabolism | Yes |  |
| 6 | cpv00053 | Ascorbate and aldarate metabolism | Yes |  |
| 7 | cpv00061 | Fatty acid biosynthesis | Yes |  |
| 8 | cpv00062 | Fatty acid elongation | Yes |  |
| 9 | cpv00071 | Fatty acid degradation | Yes |  |
| 10 | cpv00130 | Ubiquinone and other terpenoid-quinone biosynthesis | Yes |  |
| 11 | cpv00190 | Oxidative phosphorylation | Yes |  |
| 12 | cpv00230 | Purine metabolism | Yes |  |
| 13 | cpv00240 | Pyrimidine metabolism | Yes |  |
| 14 | cpv00250 | Alanine, aspartate and glutamate metabolism | Yes |  |
| 15 | cpv00260 | Glycine, serine and threonine metabolism | Yes |  |
| 16 | cpv00270 | Cysteine and methionine metabolism | Yes |  |
| 17 | cpv00310 | Lysine degradation | Yes |  |
| 18 | cpv00330 | Arginine and proline metabolism | Yes |  |
| 19 | **cpv00332** | **Carbapenem biosynthesis** | **No** | **cgd2\_2300, cgd7\_4940** |
| 20 | cpv00400 | Phenylalanine, tyrosine and tryptophan biosynthesis | Yes |  |
| 21 | cpv00440 | Phosphonate and phosphinate metabolism | Yes |  |
| 22 | cpv00450 | Selenocompound metabolism | Yes |  |
| 23 | **cpv00460** | **Cyanoamino acid metabolism** | **No** | **cgd5\_4540, cgd8\_2610** |
| 24 | cpv00480 | Glutathione metabolism | Yes |  |
| 25 | cpv00500 | Starch and sucrose metabolism | Yes |  |
| 26 | cpv00510 | N-Glycan biosynthesis | Yes |  |
| 27 | cpv00520 | Amino sugar and nucleotide sugar metabolism | Yes |  |
| 28 | cpv00561 | Glycerolipid metabolism | Yes |  |
| 29 | cpv00562 | Inositol phosphate metabolism | Yes |  |
| 30 | cpv00563 | Glycosylphosphatidylinositol (GPI)-anchor biosynthesis | Yes |  |
| 31 | cpv00564 | Glycerophospholipid metabolism | Yes |  |
| 32 | cpv00565 | Ether lipid metabolism | Yes |  |
| 33 | cpv00590 | Arachidonic acid metabolism | Yes |  |
| 34 | cpv00620 | Pyruvate metabolism | Yes |  |
| 35 | cpv00630 | Glyoxylate and dicarboxylate metabolism | Yes |  |
| 36 | cpv00640 | Propanoate metabolism | Yes |  |
| 37 | cpv00670 | One carbon pool by folate | Yes |  |
| 38 | cpv00730 | Thiamine metabolism | Yes |  |
| 39 | cpv00760 | Nicotinate and nicotinamide metabolism | Yes |  |
| 40 | cpv00770 | Pantothenate and CoA biosynthesis | Yes |  |
| 41 | cpv00790 | Folate biosynthesis | Yes |  |
| 42 | cpv00860 | Porphyrin and chlorophyll metabolism | Yes |  |
| 43 | cpv00900 | Terpenoid backbone biosynthesis | Yes |  |
| 44 | cpv00970 | Aminoacyl-tRNA biosynthesis | Yes |  |
| 45 | cpv01040 | Biosynthesis of unsaturated fatty acids | Yes |  |
| 46 | cpv01100 | Metabolic pathways | Yes |  |
| 47 | **cpv01110** | **Biosynthesis of secondary metabolites** | **No** | **cgd1\_2040, cgd1\_3020, cgd1\_3040, cgd1\_3710, cgd2\_1200, cgd2\_1770, cgd2\_210, cgd2\_2130, cgd2\_2830, cgd2\_3200, cgd2\_3260, cgd2\_3270 , cgd3\_1400, cgd3\_2100, cgd3\_2150, cgd3\_3210, cgd3\_3270 , cgd4\_1890, cgd4\_1940, cgd4\_2550, cgd4\_2790, cgd4\_4340, cgd4\_960, cgd5\_1140, cgd5\_1470, cgd5\_1960, cgd5\_3360, cgd5\_4540, cgd5\_4560, cgd6\_20, cgd6\_2450, cgd6\_3280, cgd6\_3720, cgd6\_3790, cgd6\_3800, cgd6\_4860, cgd7\_2650, cgd7\_4270, cgd7\_450, cgd7\_4940, cgd7\_910, cgd8\_1720, cgd8\_1920, cgd8\_2610, cgd8\_3680, cgd8\_790** |
| 48 | **cpv01130** | **Biosynthesis of antibiotics** | **No** | **cgd1\_2040, cgd1\_3020, cgd1\_3040, cgd1\_3710, cgd1\_3730, cgd2\_2130, cgd2\_2300, cgd2\_3200, cgd2\_3260, cgd2\_3270, cgd3\_1400, cgd3\_3270, cgd4\_1890, cgd4\_1940, cgd4\_2550, cgd4\_3310, cgd4\_810, cgd4\_960, cgd5\_1470, cgd5\_1960, cgd5\_3360, cgd5\_4560, cgd6\_2510, cgd6\_3720, cgd6\_3790, cgd6\_3800, cgd6\_70, cgd7\_1830, cgd7\_4270, cgd7\_4940, cgd7\_910, cgd8\_1720, cgd8\_1920, cgd8\_2610** |
| 49 | cpv01200 | Carbon metabolism | Yes |  |
| 50 | cpv01212 | Fatty acid metabolism | Yes |  |
| 51 | cpv01230 | Biosynthesis of amino acids | Yes |  |
| 52 | cpv02010 | ABC transporters | Yes |  |
| 53 | cpv03008 | Ribosome biogenesis in eukaryotes | Yes |  |
| 54 | cpv03010 | Ribosome | Yes |  |
| 55 | cpv03013 | RNA transport | Yes |  |
| 56 | cpv03015 | mRNA surveillance pathway | Yes |  |
| 57 | cpv03018 | RNA degradation | Yes |  |
| 58 | cpv03020 | RNA polymerase | Yes |  |
| 59 | cpv03022 | Basal transcription factors | Yes |  |
| 60 | cpv03030 | DNA replication | Yes |  |
| 61 | cpv03040 | Spliceosome | Yes |  |
| 62 | cpv03050 | Proteasome | Yes |  |
| 63 | cpv03060 | Protein export | Yes |  |
| 64 | cpv03410 | Base excision repair | Yes |  |
| 65 | cpv03420 | Nucleotide excision repair | Yes |  |
| 66 | cpv03430 | Mismatch repair | Yes |  |
| 67 | cpv03440 | Homologous recombination | Yes |  |
| 68 | cpv03450 | Non-homologous end-joining | Yes |  |
| 69 | cpv04070 | Phosphatidylinositol signaling system | Yes |  |
| 70 | cpv04120 | Ubiquitin mediated proteolysis | Yes |  |
| 71 | cpv04122 | Sulfur relay system | Yes |  |
| 72 | cpv04130 | SNARE interactions in vesicular transport | Yes |  |
| 73 | cpv04136 | Autophagy - other | Yes |  |
| 74 | cpv04141 | Protein processing in endoplasmic reticulum | Yes |  |
| 75 | cpv04144 | Endocytosis | Yes |  |
| 76 | cpv04145 | Phagosome | Yes |  |
| 77 | cpv04146 | Peroxisome | Yes |  |
| 78 | cpv04933 | AGE-RAGE signaling pathway in diabetic complications | Yes |  |

**Supplementary Table 2.** Pathway comparison data of *C. parvum* (*Cpv*) and human collected from BioCyc database

|  |  |  |  |
| --- | --- | --- | --- |
| Sl. No. | Name of *Cpv* pathways | Found in human | Genes Associated (Mentioned for only those pathways not found in human) |
| 1 | 5-aminoimidazole ribonucleotide biosynthesis II | No | cgd1\_3730 |
| 2 | N-acetylneuraminate and N-acetylmannosamine degradation I | No | NA |
| 3 | acetate formation from acetyl-CoA I | No | NA |
| 4 | acetyl-CoA fermentation to butyrate I | No | NA |
| 5 | L-arginine degradation VI (arginase 2 pathway) | No | cgd6\_3720 |
| 6 | L-ascorbate biosynthesis I (L-galactose pathway) | No | cgd8\_1920, cgd4\_960, cgd2\_1770 |
| 7 | L-asparagine biosynthesis II | No | cgd3\_110 |
| 8 | cardiolipin biosynthesis I | No | cgd3\_2940 |
| 9 | CDP-diacylglycerol biosynthesis II | No | cgd2\_210, cgd7\_450 |
| 10 | choline biosynthesis III | No | cgd8\_1150, cgd4\_2790 |
| 11 | L-citrulline degradation | No | NA |
| 12 | coenzyme A biosynthesis I (prokaryotic) | No | cgd7\_4950, cgd4\_2250, |
| 13 | ethylene biosynthesis I from methionine (plants) | No | cgd7\_2650 |
| 14 | fatty acid β-oxidation II (core pathway) (peroxisome) | No | cgd5\_3200, cgd3\_640, cgd3\_330 |
| 15 | N10-formyl-tetrahydrofolate biosynthesis (formylTHF biosynthesis I) | No | cgd4\_4460, cgd8\_2610, cgd2\_4320 |
| 16 | folate transformations II (formylTHF biosynthesis II) | No | cgd4\_4460, cgd8\_2610, cgd2\_4320 |
| 17 | GDP-mannose metabolism | No | cgd8\_1920, cgd4\_960, cgd2\_1770 |
| 18 | geranyl diphosphate biosynthesis | No | cgd4\_2550 |
| 19 | L-glutamate degradation VI (to pyruvate) | No | NA |
| 20 | glycogen degradation I | No | cgd6\_2450, cgd2\_3260, cgd2\_3270 |
| 21 | glycolysis I (from glucose 6-phosphate) | No | cgd2\_3200, cgd2\_2130, cgd1\_3020, cgd1\_3040, cgd6\_3790, cgd7\_910, cgd1\_1710, cgd5\_1960, cgd1\_2040 |
| 22 | glycolysis II (from fructose 6-phosphate) | No | cgd2\_2130, cgd1\_3020, cgd1\_3040, cgd6\_3790, cgd7\_910, cgd1\_1710, cgd5\_1960, cgd1\_2040 |
| 23 | L-histidine biosynthesis | No | NA |
| 24 | inosine-5'-phosphate biosynthesis I | No | NA |
| 25 | NAD phosphorylation and dephosphorylation | No | cgd8\_1970 |
| 26 | octane oxidation | No | cgd3\_640, cgd5\_3200 |
| 27 | L-ornithine degradation I (L-proline biosynthesis) | No | NA |
| 28 | L-ornithine degradation II (Stickland reaction) | No | cgd6\_3720 |
| 29 | phosphatidylcholine biosynthesis II | No | cgd8\_1150, cgd4\_2790 |
| 30 | phosphatidylcholine biosynthesis III | No | NA |
| 31 | phosphatidylcholine biosynthesis IV | No | NA |
| 32 | phosphatidylethanolamine biosynthesis I | No | cgd1\_1110, cgd3\_2100 |
| 33 | proline biosynthesis II | No | cgd2\_2300, cgd7\_4940, cgd6\_3720 |
| 34 | L-proline biosynthesis II (from arginine) | No | cgd6\_3720 |
| 35 | L-proline biosynthesis III | No | cgd7\_4940, cgd6\_3720 |
| 36 | PRPP biosynthesis II | No | NA |
| 37 | purine and pyrimidine metabolism | No | cgd7\_2190, cgd1\_1900, cgd8\_2810, cgd5\_4440, cgd5\_3630, cgd4\_4460, cgd5\_3360 |
| 38 | pyrimidine ribonucleotides interconversion | No | cgd4\_1940, cgd5\_1710 |
| 39 | pyruvate fermentation to ethanol II | No | cgd8\_1720 |
| 40 | pyruvate oxidation pathway | No | NA |
| 41 | salvage pathways of pyrimidine ribonucleotides | No | cgd8\_2810, cgd4\_1940, cgd1\_1900 |
| 42 | starch biosynthesis | No | cgd2\_3200, cgd2\_3260, cgd2\_3270, cgd2\_1770, cgd5\_3140, cgd6\_3280 |
| 43 | starch degradation I | No | cgd5\_2910, cgd2\_3260, cgd2\_3270, cgd6\_3800 |
| 44 | sucrose degradation III (sucrose invertase) | No | cgd2\_3260, cgd2\_3270, cgd6\_3800, cgd2\_3200 |
| 45 | sucrose degradation II (sucrose synthase) | No | cgd2\_3260, cgd2\_3270, cgd6\_3790, cgd2\_3200, cgd2\_2130, cgd1\_3020, cgd1\_3040, cgd7\_910, cgd1\_1710, cgd5\_1960, cgd1\_2040, cgd8\_1720, cgd7\_470 |
| 46 | tetrahydrofolate biosynthesis II | No | cgd4\_4460 |
| 47 | trehalose biosynthesis I | No | cgd8\_4940 |
| 48 | trehalose biosynthesis II | No | cgd8\_4940 |
| 49 | trehalose biosynthesis III | No | cgd8\_4940 |
| 50 | trehalose biosynthesis V | No | cgd5\_2910 |
| 51 | UDP-glucose conversion | No | cgd2\_3260, cgd2\_3270 |
| 52 | 5-aminoimidazole ribonucleotide biosynthesis | Yes |  |
| 53 | N-acetylglucosamine degradation I | Yes |  |
| 54 | S-methyl-5-thio-α-D-ribose 1-phosphate degradation (methionine salvage I) | Yes |  |
| 55 | trans, trans-farnesyl diphosphate biosynthesis | Yes |  |
| 56 | acetate conversion to acetyl-CoA | Yes |  |
| 57 | biotin-carboxyl carrier protein assembly | Yes |  |
| 58 | CDP-diacylglycerol biosynthesis | Yes |  |
| 59 | D-mannose degradation | Yes |  |
| 60 | dolichyl-diphosphooligosaccharide biosynthesis | Yes |  |
| 61 | ethanol degradation II | Yes |  |
| 62 | ethanol degradation IV | Yes |  |
| 63 | fatty acid β-oxidation | Yes |  |
| 64 | fatty acid activation | Yes |  |
| 65 | GDP-L-fucose biosynthesis I (from GDP-D-mannose) | Yes |  |
| 66 | GDP-mannose biosynthesis | Yes |  |
| 67 | glutathione redox reactions I | Yes |  |
| 68 | glutathione-mediated detoxification | Yes |  |
| 69 | inosine-5'-phosphate biosynthesis | Yes |  |
| 70 | lactate fermentation (reoxidation of cytosolic NADH) (pyruvate fermentation to lactate) | Yes |  |
| 71 | methionine degradation | Yes |  |
| 72 | oxidative ethanol degradation III | Yes |  |
| 73 | phosphatidylcholine biosynthesis | Yes |  |
| 74 | phosphatidylethanolamine biosynthesis II | Yes |  |
| 75 | proline biosynthesis | Yes |  |
| 76 | PRPP biosynthesis | Yes |  |
| 77 | superoxide radicals degradation | Yes |  |
| 78 | thioredoxin pathway | Yes |  |
| 79 | triacylglycerol biosynthesis | Yes |  |
| 80 | tRNA charging | Yes |  |
| 81 | UDP-N-acetyl-D-glucosamine biosynthesis II | Yes |  |
| 82 | uridine-5'-phosphate biosynthesis (UMP biosynthesis) | Yes |  |

**Supplementary Table 3.** Pathway comparison data of *Cpv* and human collected from LAMP database

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sl. No. | Name of *Cpv* pathways | Present in human | Enzymes involved (Mentioned for only those pathways not found in human) | EC Number of enzymes | Genes involved (Mentioned for only those pathways not found in human) |
| 1 | Aminosugars metabolism | Yes |  |  |  |
| 2 | CoA biosynthesis | Yes |  |  |  |
| 3 | Electron transport chain | No | Malalte:quinone oxidoreductase | 1.1.5.4 | cgd8\_380 |
| No | Alternative oxidase (ubiquinol oxidase) | 1.10.3.11 | cgd3\_3120 |
| No | NAD(P)H dehydrogenase | 1.6.5.3 | cgd7\_1900 |
| 4 | Examples of adenosylmethionine-dependent methyltransferases | Yes |  |  |  |
| 5 | Fatty acid elongation in the cytosol (FAS I system) | Yes |  |  |  |
| 6 | Fatty acid elongation via elongase pathway of ER | Yes |  |  |  |
| 7 | Glutamate metabolism | Yes |  |  |  |
| 8 | Glycine and serine metabolism | Yes |  |  |  |
| 9 | Glycolysis | Yes |  |  |  |
| 10 | Glycosylphosphatidylinositol (GPI) anchor biosynthesis | Yes |  |  |  |
| 11 | Inositol phosphate metabolism | No | Phosphoinositide-binding protein | none | cgd4\_2120, cgd7\_3980 |
| 12 | Lysine metabolism | No | Oxidoreductase | 1.1.1.- | cgd3\_2370 |
| No | Saccharopine dehydrogenase | 1.5.1.7 | cgd4\_700 |
| No | Acylaminoacyl-peptidase | 3.4.19.1 | cgd5\_4370 |
| 13 | Mannose and fructose metabolism | No | Mannitol dehydrogenase | 1.1.1.255 | cgd5\_2670 |
| 14 | Methionine metabolism | Yes |  |  |  |
| 15 | NADP synthesis | Yes |  |  |  |
| 16 | N-glycan biosynthesis | Yes |  |  |  |
| 17 | Phosphatidylcholine metabolism | No | Phospholipase C | 3.1.4.3 | cgd4\_2560 |
| No | Glycerophosphodiester phosphodiesterase | 3.1.4.46 | cgd5\_490 |
| 18 | Phosphatidylethanolamine and phosphatidylserine metabolism | No | Phospholipase C | 3.1.4.3 | cgd4\_2560 |
| No | Glycerophosphodiester phosphodiesterase | 3.1.4.46 | cgd5\_490 |
| No | MSF-1 | none | cgd5\_1390 |
| 19 | Polyamine salvage | Yes |  |  |  |
| 20 | Proline metabolism | Yes |  |  |  |
| 21 | Purine metabolism | Yes |  |  |  |
| 22 | Pyridoxal phosphate dependent enzymes | No | Alanine racemase | 5.1.1.1 | cgd5\_620 |
| 23 | Pyrimidine metabolism | Yes |  |  |  |
| 24 | Pyruvate metabolism | No | Acetoacetyl-CoA reductase | 1.1.1.36 | cgd4\_2270 |
| No | Pyruvate dehydrogenase (NADP(+)) | 1.2.1.51 | cgd4\_690 |
| No | Pyruvate decarboxylase | 4.1.1.1 | cgd7\_3120 |
| No | Phosphoenolpyruvate carboxylase | 4.1.1.31 | cgd5\_70 |
| 25 | Recycling of folate | Yes |  |  |  |
| 26 | Recycling of phospholipids | No | Phospholipase A1 | 3.1.1.32 | cgd2\_1280 |
| No | Phospholipase C | 3.1.4.3 | cgd4\_2560 |
| No | Glycerophosphodiester phosphodiesterase | 3.1.4.46 | cgd5\_490 |
| 27 | Redox metabolism | Yes |  |  |  |
| 28 | Sphingomyelin and ceramide metabolism | Yes |  |  |  |
| 29 | Starch metabolism | No | Starch synthase | 2.4.1.11 | cgd5\_3140 |
| No | Trehalose-6-phosphate synthase | 2.4.1.15 | cgd8\_4940 |
| No | Trehalose phosphatase | 3.1.3.12 | cgd8\_4940 |
| 30 | Terpenoid metabolism | No | Methyltransferase, MiaB | 2.1.1.- | cgd6\_1520 |
| No | Trans-pentaprenyltranstransferase | 2.5.1.30 | cgd7\_3730 |
| No | di-trans,poly-cis-decaprenylcistransferase | 2.5.1.31 | cgd4\_1510 |
| No | Trans-hexaprenyltranstransferase | 2.5.1.82 | cgd7\_3730 |
| No | Trans-octaprenyltranstransferase | 2.5.1.84 | cgd7\_3730 |
| 31 | Tryptophan synthesis | No | Tryptophan synthase beta subunit | 4.2.1.20 | cgd5\_4560 |
| 32 | Ubiquinone metabolism | No | FAD-dependent monooxygenase | 1.14.13.- | cgd8\_2700 |
| No | UbiB Hydroxylase | 1.14.13.- | cgd5\_2400 |
| No | Ubiquinone/menaquinone biosynthesis methyltransferase | 2.1.1.- | cgd1\_2860 |
| No | 3-demethylubiquinone-9 3-O-methyltransferase | 2.1.1.64 | cgd2\_2830, cgd5\_4330 |
| No | p-hydroxybenzoate-polyprenyltransferase | 2.5.1.- | cgd3\_2460 |
| 33 | Vacuolar-ATPase | Yes |  |  |  |

**Supplementary Table 4.** Eleven non-homologous proteins of *C. parvum* when blastp search against human host proteome at an e-value of 0.005, along with essentiality assessment result when manual blastp search was performed against *Cryptosporidium* taxon at an e value of 1e-05.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sl. No. | Gene ID | Protein Name | UNIPROTKB ID | Essentiality (Other related species apart from *C. parvum* in which protein is found when blastp was performed at e value of 1e-05) |
| 1 | cgd2\_2130 | Probable ATP-dependent 6-phosphofructokinase | Q5CTQ1 | *C. hominis, C. muris, C. ubiquitum, C. andersoni* |
| 2 | cgd3\_1400 | Pyrophosphate-fructose 6-phosphate 1-phosphotransferase | Q5CUU9 | *C. hominis, C. muris, C. ubiquitum, C. andersoni* |
| 3 | cgd3\_2940 | Probable phosphatidylserine/phosphatidylglycerophosphate/cardiolipi n synthase, 2x SMART\_PLDc domains, possible bacterial origin | Q5CUG3 | *C. hominis, C. muris, C. ubiquitum, C. andersoni* |
| 4 | cgd3\_3120 | AOX1,alternative oxidase, possible fungal or bacterial origin, 2 transmembrane regions | Q5CUE8 | *C. hominis, C. muris, C. andersoni* |
| 5 | cgd5\_70 | Phosphoenolpyruvate carboxylase | Q5CS64 | *C. hominis, C. muris, C. ubiquitum, C. andersoni* |
| 6 | cgd5\_4440 | Thymidine kinase | Q5CQ21 | *C. hominis, C. muris, C. ubiquitum, C. andersoni* |
| 7 | cgd5\_4540 | Asparagine synthetase A (AsnA) like protein | Q5CPD9 | *C. hominis, C. ubiquitum* |
| 8 | cgd5\_4560 | Tryptophan synthase trpB of possible bacterial origin | Q5CPD7 | *C. hominis, C. ubiquitum* |
| 9 | cgd7\_1830 | Secreted UDP-N-acetylglucosamine pyrophosphorylase family protein, signal peptide | Q5CYM0 | *C. hominis, C. muris, C. ubiquitum, C. andersoni* |
| 10 | cgd8\_380 | Possible oxidase or dehydrogenase | Q5CPY0 | *C. hominis, C. muris, C. ubiquitum, C. andersoni* |
| 11 | cgd8\_4940 | Trehalose-6-phosphate synthase of likely plant origin | Q5CV93 | *C. hominis, C. muris, C. ubiquitum, C. andersoni* |

**Supplementary Table 5.** Top alleles selected in MHC class I and MHC class II as per their binding score for cgd\_1400 and cgd3\_2940 with ProPred I (with a cut-off of 90%) and ProPred (with a cut off of 70%) respectively.

|  |  |  |
| --- | --- | --- |
| ProPred I: MHC Class I Alleles for cgd3\_1400 | | |
| Allele Name | Binding Score | Percentage of highest on Log Scale (in %) |
| MHC- Kd | 5760 | 104.46 |
| HLA- A20 Cattle | 4000 | 100.00 |
| HLA- B\*51 | 145.83 | 92.89 |
| HLA- B\*5301 | 139.43 | 92.64 |
| ProPred I: MHC Class II Alleles for cgd3\_1400 | | |
| DRB1\_1304 | 6.8000 | 75.56 |
| ProPred I: MHC Class I Alleles for cgd3\_2940 | | |
| MHC- Kd | 6912 | 106.65 |
| HLA-B\*2705 | 30000 | 100 |
| HLA-B\*51 | 147.15 | 93.73 |
| HLA-B\*5301 | 139.22 | 92.50 |
| ProPred: MHC Class II Alleles for cgd3\_2940 | | |
| DRB1\_1120 | 6.5 | 73.86 |
| DRB1\_1302 | 6.5 | 73.86 |
| DRB1\_1304 | 6.6 | 73.33 |
| DRB1\_1304 | 6.6 | 73.33 |
| DRB1\_1327 | 6.4 | 72.73 |
| DRB1\_1328 | 6.4 | 72.73 |
| DRB1\_1301 | 6.4 | 72.73 |

**Supplementary Table 6**. Additional Parameters used on the proposed vaccine candidate proteins cgd3\_1400 and cgd3\_2940

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| UniProtKB ID | Gene ID | Length | SignalP (signal peptide prediction) | TMHMM (helix prediction) | VaxiJen (antigen probability) | IgPred (Ig isotype prediction) | AlgPred (allergen prediction) |
| [Q5CUU9](http://www.uniprot.org/uniprot/Q5CUU9) | cgd3\_1400 | 1327 | No | 0 | 0.485 (Non-Antigen) | 1.413 (IgG) | -0.603 (Non-Allergen) |
| Q5CUG3 | cgd3\_2940 | 476 | No | 0 | 0.529 (Antigen) | 1.497 (IgG) | 0.348 (Allergen) |

**Supplementary Table 7.** Comparison of allele-peptide binding score distribution of cgd3\_1400 and cgd3\_2940

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Proteins | ProPred1: MHC class I alleles | | ProPred: MHC class II alleles | |
| Peptides at 90-100% binding score with MHC alleles | Peptides at 100% binding score with MHC alleles | Peptides at 70-80% binding score with MHC alleles | Peptides at 80-90% binding score with MHC alleles |
| cgd3\_1400 | 2 | 2 | 1 | 0 |
| cgd3\_2940 | 2 | 2 | 7 | 0 |

**Supplementary Table 8.** Top threeB-cell epitopes and T-cell epitopes on cgd3\_1400 predicted by ABCpred and CTLPred respectively

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| B-Cell Epitopes | | | | |
| ABCPred Rank | Start Position | Sequence | Score | DISOPRED prediction |
| 1 | 850 | HQMIETCIGFDSVTKS | 0.97 | Not disordered |
| 2 | 471 | QRSGYMAVVRKLSLTP | 0.96 | Not disordered |
| 3 | 1205 | PGPIQYFNLFKNLFNR | 0.95 | Not disordered |
| T- Cell Epitopes | | | | |
| CTLPred Rank | Start Position | Sequence | Score (ANN/SVM) | DISOPRED prediction |
| 1 | 1202 | YANPGPIQY | 0.67/1.3020669 | Not disordered |
| 2 | 731 | SLAFFENSL | 0.82/1.1142642 | Not disordered |
| 3 | 1000 | SLSRYEKKV | 0.97/0.90808523 | Not disordered |

**Supplementary Table 9.** Docking result for cgd7\_1830 protein & ligands retrieved from various databases with reference to active sites prediction by 3DLigandSite server & lead compound by COFACTOR server (Top ten compounds arranged in decreasing order of Rank Score)

|  |  |  |
| --- | --- | --- |
| Sl. No. | Compound ID | Rank Score (Kcal/mol) |
| 1 | PubChem\_57371104 | -18.966 |
| 2 | PubChem\_92021883 | -17.653 |
| 3 | PubChem\_22960366 | -15.938 |
| 4 | PubChem\_46232346 | -15.732 |
| 5 | PubChem\_42627746 | -15.602 |
| 6 | PubChem\_22062379 | -15.197 |
| 7 | PubChem\_11852357 | -15.187 |
| 8 | ChEMBL\_90669233 | -14.284 |
| 9 | ChEMBL\_9872620 | -13.962 |
| 10 | PubChem\_9830339 | -13.713 |