

Single-nucleotide variations associated with *Mycobacterium tuberculosis* KwaZulu-Natal strains

SARBASHIS DAS, RAGOTHAMAN M YENNAMALLI, ANCHAL VISHNOI, PARUL GUPTA and ALOK BHATTACHARYA

J. Biosci. 34(3), September 2009, 397–404 © Indian Academy of Sciences

Supplementary tables

Supplementary table S1. Unique single-nucleotide variations (SNVs) of KZN1435 (multidrug-resistant [MDR]), KZN4207 (drug-sensitive [DS]) and KZN605 (extensively drug-resistant [XDR]) strains of *Mycobacterium tuberculosis*

KZN1435 (MDR) unique SNVs

Position	SNP	Gene	COG	Gene Description
J - Translation, ribosomal structure and biogenesis				
2144616	g-a	Rv1897c	COG1490J	D-tyrosyl-tRNA deacylase
3877420	a-g	Rv3456c	COG0203J	50S ribosomal protein L17
4042815	a-g	Rv3598c	COG1190J	lysyl-tRNA synthetase
K - Transcription				
86602	c-g	Rv0078	COG1309K	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN
761108	g-t	Rv0667	COG0085K	DNA-directed RNA polymerase subunit beta
761265	a-g	Rv0667	COG0085K	DNA-directed RNA polymerase subunit beta
3833050	c-t	Rv3414c	COG1595K	RNA polymerase sigma factor SigD
L - DNA replication, recombination and repair				
2250	c-t	Rv0002	COG0592L	DNA polymerase III subunit beta
2526	a-g	Rv0002	COG0592L	DNA polymerase III subunit beta
60773	g-a	Rv0058	COG0305L,COG1372L	replicative DNA helicase
61049	g-t	Rv0058	COG0305L,COG1372L	replicative DNA helicase
62174	a-c	Rv0058	COG0305L,COG1372L	replicative DNA helicase
62295	g-n	Rv0058	COG0305L,COG1372L	replicative DNA helicase
62301	g-n	Rv0058	COG0305L,COG1372L	replicative DNA helicase
62906	c-a	Rv0058	COG0305L,COG1372L	replicative DNA helicase
4400070	g-c	Rv3910	COG0515RTKL,COG0728R	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN
D - Cell division and chromosome partitioning				
3864994	t-c	Rv3447c	COG1674D	PROBABLE CONSERVED MEMBRANE PROTEIN
3867395	t-a	Rv3447c	COG1674D	PROBABLE CONSERVED MEMBRANE PROTEIN
4379679	c-g	Rv3894c	COG1674D	POSSIBLE CONSERVED MEMBRANE PROTEIN
O - Posttranslational modification, protein turnover,chaperones				
151284	c-t	Rv0125	COG0265O	PROBABLE SERINE PROTEASE PEPA (SERINE PROTEINASE) (MTB32A)
2376079	a-n	Rv2115c	COG0464O	Probable ATPase
4038286	g-a	Rv3596c	COG0542O	PROBABLE ATP-DEPENDENT PROTEASE ATP-BINDING SUBUNIT CLPC1

M - Cell envelope biogenesis

3796788 c-t Rv3382c COG0761IM
 3856223 g-a Rv3436c COG0449M

PROBABLE LYTB-RELATED PROTEIN LYTB1
 D-fructose-6-phosphate amidotransferase

P - Inorganic ion transport and metabolism

127014 c-a Rv0107c COG0474P
 127134 a-g Rv0107c COG0474P
 127290 g-a Rv0107c COG0474P
 128415 c-t Rv0107c COG0474P
 129147 g-c Rv0107c COG0474P
 2155167 c-g Rv1908c COG0376P

PROBABLE CATION-TRANSPORTER ATPASE I CTPI
 PROBABLE CATION-TRANSPORTER ATPASE I CTPI
 PROBABLE CATION-TRANSPORTER ATPASE I CTPI
 PROBABLE CATION-TRANSPORTER ATPASE I CTPI
 PROBABLE CATION-TRANSPORTER ATPASE I CTPI
 CATALASE-PEROXIDASE-PEROXYNITRITASE T
 KATG

T - Singal Transduction mechanisms

3776705 c-t Rv3365c COG0642T

hypothetical protein

C - Energy production and conversion

72735 g-a Rv0066c COG2838C

PROBABLE ISOCITRATE DEHYDROGENASE [NADP]
 ICD2 (OXALOSUCCINATE DECARBOXYLASE) (IDH)
 (NADP+-SPECIFIC ICDH) (IDP)

G - Carbohydrate transport and metabolism

152409 g-a Rv0126 COG0366G
 153141 c-a Rv0126 COG0366G
 155087 c-g Rv0127 COG3281G

TREHALOSE SYNTHASE TRES
 TREHALOSE SYNTHASE TRES
 hypothetical protein

E - Amino acid transport and metabolism

76814 c-g Rv0069c COG1760E
 84162 g-a Rv0075 COG1168E
 138059 g-a Rv0114 COG0241E

PROBABLE L-SERINE DEHYDRATASE SDAA (L-SERINE DEAMINASE) (SDH) (L-SD)
 PROBABLE AMINOTRANSFERASE
 POSSIBLE D-ALPHA,BETA-D-HEPTOSE-1,7-BIPHOSPHATE PHOSPHATASE GMHB (D-GLYCERO-D-MANNO-HEPTOSE 7-PHOSPHATE KINASE)
 CONSERVED EXPORTED PROTEIN
 aspartate kinase

F - Nucleotide transport and mechnism

3829769 t-c Rv3410c COG0516F

inositol-5-monophosphate dehydrogenase

H - Coenzyme metabolism

142523 c-a Rv0118c COG0028EH
 1777212 c-g Rv1569 COG0156H
 2193247 g-a Rv1940 COG0108H,COG0807H
 3714756 a-c Rv3329 COG0161H
 3798094 a-c Rv3383c COG0142H

putative oxalyl-CoA decarboxylase
 8-amino-7-oxononanoate synthase
 Probable Riboflavin biosynthesis protein ribA1 (GTP cyclohydrolase II)
 hypothetical protein
 POSSIBLE POLYPRENYL SYNTHETASE IDSB (POLYPRENYL TRANSFERASE) (POLYPRENYL DIPHOSPHATE SYNTHASE)

I - Lipid metabolism

3191026 g-a Rv2881c COG0575I

PROBABLE INTEGRAL MEMBRANE PHOSPHATIDATE CYTIDYLYLTRANSFERASE CDSA (CDP-DIGLYCERIDE SYNTHETASE) (CDP-DIGLYCERIDE PYROPHOSPHORYLASE) (CDP-DIACYLGLYCEROL SYNTHASE) (CDS) (CTP: PHOSPHATIDATE CYTIDYLYLTRANSFERASE) (CDP-DAG SYNTHASE) (CDP-DG SYNTHETASE)

3952799	g-a	Rv3516	COG1024I	enoyl-CoA hydratase
4011949	g-a	Rv3570c	COG1960I	POSSIBLE OXIDOREDUCTASE
Q - Secondary metabolites biosynthesis, transport and catabolism				
37816	g-a	Rv0035	COG0318IQ	PROBABLE FATTY-ACID-CoA LIGASE FADD34 (FATTY-ACID-CoA SYNTHETASE) (FATTY-ACID-CoA SYNTHASE)
163413	c-g	Rv0136	COG2124Q	PROBABLE CYTOCHROME P450 138 CYP138
163436	g-t	Rv0136	COG2124Q	PROBABLE CYTOCHROME P450 138 CYP138
1001603	g-a	Rv0897c	COG1233Q	PROBABLE OXIDOREDUCTASE
2288846	c-g	Rv2043c	COG1335Q	PYRAZINAMIDASE/NICOTINAMIDAS PNCA (PZase)
2296041	g-c	Rv2048c	COG3321Q	Probable polyketide synthase pks12
2674577	g-c	Rv2383c	COG1020Q,COG3208Q	PHENYLOXAZOLINE SYNTHASE MBTB (PHENYLOXAZOLINE SYNTHETASE)
3817116	c-a	Rv3399	COG3315Q	hypothetical protein
3825559	c-g	Rv3406	COG2175Q	PROBABLE DIOXYGENASE
4293071	g-a	Rv3824c	COG1020Q	PROBABLE CONSERVED POLYKETIDE SYNTHASE ASSOCIATED PROTEIN PAPA1
R - General function prediction only				
39949	g-a	Rv0037c	COG0477GEPR	PROBABLE CONSERVED INTEGRAL MEMBRANE PROTEIN
133533	t-c	Rv0110	COG0705R	PROBABLE CONSERVED INTEGRAL MEMBRANE PROTEIN
133673	g-c	Rv0110	COG0705R	PROBABLE CONSERVED INTEGRAL MEMBRANE PROTEIN
161991	t-c	Rv0134	COG0596R	POSSIBLE EPOXIDE HYDROLASE EPHF (EPOXIDE HYDRATASE) (ARENE-OXIDE HYDRATASE)
1944401	t-c	Rv1716	COG1878R	hypothetical protein
3718356	c-t	Rv3331	COG0477GEPR	PROBABLE SUGAR-TRANSPORT INTEGRAL MEMBRANE PROTEIN SUGI
3838870	a-g	Rv3420c	COG0456R	PROBABLE RIBOSOMAL-PROTEIN-ALANINE ACETYLTRANSFERASE RIMI (ACETYLATED ENZYME FOR N-TERMINAL OF RIBOSOMAL PROTEIN S18)
S - Function unknown				
1995470	g-a	Rv1762c	COG0393S	hypothetical protein
2958387	c-n	Rv2631	COG1690S	hypothetical protein
3823158	a-t	Rv3403c	COG4529S	hypothetical protein
KZN4207 (DS) unique SNVs				
Position	SNP	Gene	COG	Gene Description
J - Translation, ribosomal structure and biogenesis				
1917971	a-g	Rv1694	COG1189J	CYTOTOXIN HAEMOLYSIN HOMOLOGUE TLYA
K - Transcription				
791485	g-a	Rv0691c	COG1309K	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN
3903529	g-c	Rv3484	COG1316K	POSSIBLE CONSERVED PROTEIN CPSA
L - DNA replication, recombination and repair				
9039	a-g	Rv0006	COG0188L	DNA gyrase subunit A
9072	c-g	Rv0006	COG0188L	DNA gyrase subunit A

9295	t-g	Rv0006	COG0188L	DNA gyrase subunit A
D - Cell division and chromosome partitioning				
3177967	g-a	Rv2866	COG2026JD	hypothetical protein
4348126	c-t	Rv3870	COG1674D	POSSIBLE CONSERVED TRANSMEMBRANE PROTEIN
O - Posttranslational modification, protein turnover, chaperones				
235007	g-t	Rv0198c	COG3590O	PROBABLE ZINC METALLOPROTEASE
235040	c-g	Rv0198c	COG3590O	PROBABLE ZINC METALLOPROTEASE
235264	c-g	Rv0198c	COG3590O	PROBABLE ZINC METALLOPROTEASE
236210	t-g	Rv0198c	COG3590O	PROBABLE ZINC METALLOPROTEASE
302557	t-c	Rv0251c	COG0071O	HEAT SHOCK PROTEIN HSP (HEAT-STRESS-INDUCED RIBOSOME-BINDING PROTEIN A)
M - Cell envelope biogenesis				
4233298	g-a	Rv3786c	COG0463M, COG0739M	hypothetical protein
4407903	g-a	Rv3919c	COG0357M	glucose-inhibited division protein B
4408155	a-c	Rv3919c	COG0357M	glucose-inhibited division protein B
P - Inorganic ion transport and metabolism				
122108	a-g	Rv0103c	COG2217P	PROBABLE CATION-TRANSPORTER P-TYPE ATPASE B CTPB
127179	c-g	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
127212	t-c	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
127436	c-g	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
129114	g-t	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
C - Energy production and conversion				
73354	c-a	Rv0066c	COG2838C	PROBABLE ISOCITRATE DEHYDROGENASE [NADP] ICD2 (OXALOSUCCINATE DECARBOXYLASE) (IDH) (NADP+-SPECIFIC ICDH) (IDP)
73387	g-c	Rv0066c	COG2838C	PROBABLE ISOCITRATE DEHYDROGENASE [NADP] ICD2 (OXALOSUCCINATE DECARBOXYLASE) (IDH) (NADP+-SPECIFIC ICDH) (IDP)
73611	g-a	Rv0066c	COG2838C	PROBABLE ISOCITRATE DEHYDROGENASE [NADP] ICD2 (OXALOSUCCINATE DECARBOXYLASE) (IDH) (NADP+-SPECIFIC ICDH) (IDP)
90143	a-g	Rv0082	COG3260C	PROBABLE OXIDOREDUCTASE
279762	t-c	Rv0234c	COG1012C	succinic semialdehyde dehydrogenase
280494	c-a	Rv0234c	COG1012C	succinic semialdehyde dehydrogenase
4315383	t-c	Rv3842c	COG0584C	PROBABLE GLYCEROPHOSPHORYL DIESTER PHOSPHODIESTERASE GLPQ1 (GLYCEROPHOSPHODIESTER PHOSPHODIESTERASE)
G - Carbohydrate transport and metabolism				
153035	a-g	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
154787	g-a	Rv0127	COG3281G	hypothetical protein
154820	a-c	Rv0127	COG3281G	hypothetical protein
155044	g-c	Rv0127	COG3281G	hypothetical protein
2812702	c-t	Rv2498c	COG2301G	PROBABLE CITRATE (PRO-3S)-LYASE (BETA SUBUNIT) CITE (CITRASE) (CITRATASE) (CITRITASE) (CITRIDESMOLASE) (CITRASE ALDOLASE)

F - Nucleotide transport and mechanism

278816	t-g	Rv0233	COG0208F	ribonucleotide-diphosphate reductase subunit beta
--------	-----	--------	----------	---

H - Coenzyme metabolism

307409	t-c	Rv0255c	COG1492H	cobyric acid synthase
1591661	c-t	Rv1415	COG0108H, COG0807H	bifunctional 3,4-dihydroxy-2-butanone 4-phosphate synthase/GTP cyclohydrolase II protein

I - Lipid metabolism

293335	g-t	Rv0243	COG0183I	acetyl-CoA acetyltransferase
294899	c-g	Rv0244c	COG1960I	PROBABLE ACYL-CoA DEHYDROGENASE FADE5
294932	t-a	Rv0244c	COG1960I	PROBABLE ACYL-CoA DEHYDROGENASE FADE5
295156	t-a	Rv0244c	COG1960I	PROBABLE ACYL-CoA DEHYDROGENASE FADE5
4255921	a-g	Rv3799c	COG4799I	PROBABLE PROPIONYL-CoA CARBOXYLASE BETA CHAIN 4 ACCD4 (PCCASE) (PROPANOYL-COA: CARBON DIOXIDE LIGASE)

Q - Secondary metabolites biosynthesis, transport and catabolism

37817	t-a	Rv0035	COG0318IQ	PROBABLE FATTY-ACID-CoA LIGASE FADD34 (FATTY-ACID-CoA SYNTHETASE) (FATTY-ACID-CoA SYNTHASE)
38041	g-a	Rv0035	COG0318IQ	PROBABLE FATTY-ACID-CoA LIGASE FADD34 (FATTY-ACID-CoA SYNTHETASE) (FATTY-ACID-CoA SYNTHASE)
82998	g-t	Rv0074	COG1228Q	hypothetical protein
171900	c-a	Rv0145	COG3315Q	hypothetical protein
171933	t-g	Rv0145	COG3315Q	hypothetical protein
172157	a-g	Rv0145	COG3315Q	hypothetical protein
173103	g-t	Rv0146	COG3315Q	hypothetical protein
2300545	a-t	Rv2048c	COG3321Q	Probable polyketide synthase pks12
2674094	g-a	Rv2383c	COG1020Q,COG3208Q	PHENYLOXAZOLINE SYNTHASE MBTB (PHENYLOXAZOLINE SYNTHETASE)
3973953	c-t	Rv3535c	COG4569Q	acetaldehyde dehydrogenase

R - General function prediction only

156722	c-a	Rv0129c	COG0627R	SECRETED ANTIGEN 85-C FBPC (85C) (ANTIGEN 85 COMPLEX C) (AG58C) (MYCOLYL TRANSFERASE 85C) (FIBRONECTIN-BINDING PROTEIN C)
291069	c-t	Rv0242c	COG1028IQR	3-ketoacyl-(acyl-carrier-protein) reductase
3938659	a-g	Rv3510c	COG2159R	hypothetical protein
4382053	t-c	Rv3896c	COG3953R	hypothetical protein

S - Function unknown

155990	g-a	Rv0128	COG3619S	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN
223941	t-c	Rv0192	COG1376S	hypothetical protein
242665	g-a	Rv0204c	COG0392S	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN
247517	a-c	Rv0207c	COG1432S	hypothetical protein
4174563	t-c	Rv3727	COG3349S	POSSIBLE OXIDOREDUCTASE
4187816	a-g	Rv3737	COG2966S,COG3610S	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN

KZN605 (XDR) unique SNVs

Position	SNP	Gene	COG	Gene Description
J - Translation, ribosomal structure and biogenesis				
3367764	g-a	Rv3009c	COG0064J	aspartyl/glutamyl-tRNA amidotransferase subunit B
K - Transcription				
162768	c-a	Rv0135c	COG1309K	POSSIBLE TRANSCRIPTIONAL REGULATORY PROTEIN
761109	a-g	Rv0667	COG0085K	DNA-directed RNA polymerase subunit beta
761160	t-c	Rv0667	COG0085K	DNA-directed RNA polymerase subunit beta
763122	t-c	Rv0667	COG0085K	DNA-directed RNA polymerase subunit beta
1526818	c-a	Rv1358	COG2114T,COG2771K,COG3903R	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN
3630177	c-a	Rv3249c	COG1309K	POSSIBLE TRANSCRIPTIONAL REGULATORY PROTEIN (PROBABLY TETR-FAMILY)
L - DNA replication, recombination and repair				
4012	t-c	Rv0003	COG1195L	recombination protein F
7361	g-c	Rv0006	COG0188L	DNA gyrase subunit A
7569	c-t	Rv0006	COG0188L	DNA gyrase subunit A
7584	g-c	Rv0006	COG0188L	DNA gyrase subunit A
890548	g-a	Rv0797	COG3547L	IS1547 transposase
890717	a-c	Rv0797	COG3547L	IS1547 transposase
1170403	c-a	Rv1047	COG3328L	PROBABLE TRANSPOSASE
1341623	g-t	Rv1199c	COG3328L	POSSIBLE TRANSPOSASE
1342062	t-c	Rv1199c	COG3328L	POSSIBLE TRANSPOSASE
1342580	t-c	Rv1199c	COG3328L	POSSIBLE TRANSPOSASE
1579137	g-n	Rv1402	COG1198L	primosome assembly protein PriA
2361603	c-g	Rv2101	COG0553KL	PROBABLE HELICASE HELZ
2361622	a-c	Rv2101	COG0553KL	PROBABLE HELICASE HELZ
2362040	c-a	Rv2101	COG0553KL	PROBABLE HELICASE HELZ
2439408	c-t	Rv2177c	COG3547L	POSSIBLE TRANSPOSASE
2439518	g-a	Rv2177c	COG3547L	POSSIBLE TRANSPOSASE
2828821	g-t	Rv2512c	COG3328L	IS1081 transposase
2829260	t-c	Rv2512c	COG3328L	IS1081 transposase
2829778	t-c	Rv2512c	COG3328L	IS1081 transposase
3117999	a-g	Rv2812	COG2801L	PROBABLE TRANSPOSASE
3381640	g-t	Rv3023c	COG3328L	PROBABLE TRANSPOSASE
3579494	c-g	Rv3202c	COG0210L,COG2887L	POSSIBLE ATP-DEPENDENT DNA HELICASE
3579497	a-t	Rv3202c	COG0210L,COG2887L	POSSIBLE ATP-DEPENDENT DNA HELICASE
3801266	g-a	Rv3387	COG3039L	POSSIBLE TRANSPOSASE
4077162	c-a	Rv3638	COG1484L	transposase
4182510	c-g	Rv3731	COG1793L	ATP-dependent DNA ligase
4182733	a-g	Rv3731	COG1793L	ATP-dependent DNA ligase
O - Posttranslational modification, protein turnover,chaperones				
151420	a-c	Rv0125	COG0265O	PROBABLE SERINE PROTEASE PEPA (SERINE PROTEINASE) (MTB32A)

M - Cell envelope biogenesis

1361109	c-n	Rv1217c	COG3559M	PROBABLE TETRONASIN-TRANSPORT INTEGRAL MEMBRANE PROTEIN ABC TRANSPORTER
2939372	g-c	Rv2611c	COG1560M	lipid A biosynthesis lauroyl acyltransferase
3175701	t-c	Rv2864c	COG0768M	POSSIBLE PENICILLIN-BINDING LIPOPROTEIN
4408110	a-c	Rv3919c	COG0357M	glucose-inhibited division protein B

P - Inorganic ion transport and metabolism

94151	a-g	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94274	g-c	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94316	t-g	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94372	c-t	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94395	g-a	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94414	a-g	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94422	g-a	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94428	a-g	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94464	t-c	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94477	g-t	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94484	t-c	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94496	c-g	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94557	c-a	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94562	c-g	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94621	g-t	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94643	a-c	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
95123	t-a	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
99164	g-t	Rv0090	COG0601EP,COG3305S	POSSIBLE MEMBRANE PROTEIN
100708	c-g	Rv0092	COG2217P	PROBABLE CATION TRANSPORTER P-TYPE ATPASE A CTPA
129149	c-g	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
129579	t-g	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
129750	t-g	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
129829	c-a	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
129834	c-a	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
129848	c-g	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
129862	g-c	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
129892	g-c	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
129902	g-a	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
129911	g-c	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
129922	c-a	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
129991	t-c	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
130077	t-a	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
756756	c-t	Rv0663	COG3119P	POSSIBLE ARYLSULFATASE ATSD (ARYL-SULFATE SULPHOHYDROLASE) (ARYLSULPHATASE)
3483698	g-a	Rv3116	COG0476H,COG0607P	PROBABLE MOLYBDENUM COFACTOR BIOSYNTHESIS PROTEIN MOEB2 (MPT-SYNTHASE SULFURYLAZE) (MOLYBDOPTERIN SYNTHASE SULPHURYLAZE)

C - Energy production and conversion

93715	t-g	Rv0085	COG4237C	POSSIBLE HYDROGENASE HYCP
93718	t-g	Rv0085	COG4237C	POSSIBLE HYDROGENASE HYCP
1292101	a-g	Rv1162	COG1140C	PROBABLE RESPIRATORY NITRATE REDUCTASE (BETA CHAIN) NARH

T - Singal Transduction mechanisms

24124	g-a	Rv0020c	COG1716T	hypothetical protein
3596543	c-t	Rv3220c	COG3920T	PROBABLE TWO COMPONENT SENSOR KINASE

G - Carbohydrate transport and metabolism

152673	c-t	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
152839	t-g	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
152895	t-c	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
152918	g-t	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
152923	c-a	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
152931	c-t	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
152937	c-g	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
152945	a-t	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
152951	a-c	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
152987	a-g	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
152991	c-g	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
153000	t-g	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
153007	g-t	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
153027	t-g	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
153080	a-c	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
153085	c-a	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
153144	c-g	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
153166	a-g	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
153600	g-t	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
154866	c-t	Rv0127	COG3281G	hypothetical protein
1630147	a-c	Rv1449c	COG0021G	transketolase
2470148	t-c	Rv2205c	COG1929G	hypothetical protein

E - Amino acid transport and metabolism

2246031	t-c	Rv2000	COG0493ER,COG0665E	hypothetical protein
---------	-----	--------	--------------------	----------------------

F - Nucleotide transport and mechnism

99781	c-t	Rv0091	COG0775F	PROBABLE BIFUNCTIONAL MTA/SAH NUCLEOSIDASE MTN: 5'-METHYLTHIOADENOSINE NUCLEOSIDASE (METHYLTHIOADENOSINE METHYLTHIORIBOHYDROLASE) + S-ADENOSYLHOMOCYSTEINE NUCLEOSIDASE (S-ADENOSYL-L-HOMOCYSTEINE HOMOCYSTEINYLRIBOHYDROLASE)
-------	-----	--------	----------	--

H - Coenzyme metabolism

2509139	g-c	Rv2236c	COG1270H	cobalamin biosynthesis protein
---------	-----	---------	----------	--------------------------------

I - Lipid metabolism

133976	c-a	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134032	g-c	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE

134039	a-c	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134042	g-t	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134055	t-c	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134060	g-a	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134064	t-a	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134074	a-c	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134084	c-g	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134088	t-a	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134124	g-c	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134137	c-a	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134143	c-t	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134148	t-g	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134217	a-c	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134222	a-g	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134281	c-t	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134737	g-c	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
135215	c-g	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
136003	t-g	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
159231	t-c	Rv0131c	COG1960I	PROBABLE ACYL-CoA DEHYDROGENASE FADE1
294011	t-g	Rv0244c	COG1960I	PROBABLE ACYL-CoA DEHYDROGENASE FADE5
2786951	a-g	Rv2482c	COG2937I	glycerol-3-phosphate acyltransferase
3606367	t-g	Rv3229c	COG3239I	POSSIBLE LINOLEOYL-CoA DESATURASE (DELTA(6)-DESATURASE)
3953252	t-c	Rv3516	COG1024I	enoyl-CoA hydratase
Q - Secondary metabolites biosynthesis, transport and catabolism				
38710	g-c	Rv0035	COG0318IQ	PROBABLE FATTY-ACID-CoA LIGASE FADD34 (FATTY-ACID-CoA SYNTHETASE) (FATTY-ACID-CoA SYNTHASE)
922003	c-t	Rv0830	COG3315Q	hypothetical protein
1327889	g-a	Rv1186c	COG2508TQ	hypothetical protein
2300236	a-g	Rv2048c	COG3321Q	Probable polyketide synthase pks12
2300665	g-a	Rv2048c	COG3321Q	Probable polyketide synthase pks12
2300668	g-c	Rv2048c	COG3321Q	Probable polyketide synthase pks12
2300689	g-a	Rv2048c	COG3321Q	Probable polyketide synthase pks12
2300698	g-a	Rv2048c	COG3321Q	Probable polyketide synthase pks12
2302032	g-a	Rv2048c	COG3321Q	Probable polyketide synthase pks12
2674131	-	Rv2383c	COG1020Q,COG3208Q	PHENYLOXAZOLINE SYNTHASE MBTB (PHENYLOXAZOLINE SYNTHETASE)
2676709	-	Rv2384	COG1021Q	BIFUNCTIONAL ENZYME MBTA: SALICYL-AMP LIGASE (SAL-AMP LIGASE) + SALICYL-S-ArCP SYNTHETASE
4256911	t-c	Rv3800c	COG3319Q,COG3321Q	POLYKETIDE SYNTHASE PKS13
R - General function prediction only				
133375	t-g	Rv0110	COG0705R	PROBABLE CONSERVED INTEGRAL MEMBRANE PROTEIN
138824	c-g	Rv0115	COG2605R	POSSIBLE D-ALPHA-D-HEPTOSE-7-PHOSPHATE KINASE HDDA

1586248	c-t	Rv1410c	COG0477GEPR	AMINOGLYCOSIDES/TETRACYCLINE-TRANSPORT INTEGRAL MEMBRANE PROTEIN
2269779	t-c	Rv2024c	COG4889R	hypothetical protein
2808476	c-t	Rv2494	COG1848R	hypothetical protein
4306422	c-g	Rv3832c	COG0500QR	hypothetical protein
4306591	g-c	Rv3832c	COG0500QR	hypothetical protein
4306601	c-g	Rv3832c	COG0500QR	hypothetical protein

S - Function unknown

14250	g-a	Rv0012	COG3879S	PROBABLE CONSERVED MEMBRANE PROTEIN
140368	t-g	Rv0116c	COG1376S	POSSIBLE CONSERVED MEMBRANE PROTEIN
1093405	a-g	Rv0978c	COG3391S	PE-PGRS FAMILY PROTEIN
1095230	g-n	Rv0980c	COG3391S	PE-PGRS FAMILY PROTEIN
1095354	g-n	Rv0980c	COG3391S	PE-PGRS FAMILY PROTEIN
1095428	g-n	Rv0980c	COG3391S	PE-PGRS FAMILY PROTEIN
1095479	g-n	Rv0980c	COG3391S	PE-PGRS FAMILY PROTEIN
1096397	c-t	Rv0980c	COG3391S	PE-PGRS FAMILY PROTEIN
3135911	g-c	Rv2828c	COG4293S	hypothetical protein

Position	SNP	Gene	COG	Gene Description
----------	-----	------	-----	------------------

J - Translation, ribosomal structure and biogenesis

2144616	g-a	Rv1897c	COG1490J	D-tyrosyl-tRNA deacylase
3877420	a-g	Rv3456c	COG0203J	50S ribosomal protein L17
4042815	a-g	Rv3598c	COG1190J	lysyl-tRNA synthetase

K - Transcription

86602	c-g	Rv0078	COG1309K	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN
761108	g-t	Rv0667	COG0085K	DNA-directed RNA polymerase subunit beta
761265	a-g	Rv0667	COG0085K	DNA-directed RNA polymerase subunit beta
3833050	c-t	Rv3414c	COG1595K	RNA polymerase sigma factor SigD

L - DNA replication, recombination and repair

2250	c-t	Rv0002	COG0592L	DNA polymerase III subunit beta
2526	a-g	Rv0002	COG0592L	DNA polymerase III subunit beta
60773	g-a	Rv0058	COG0305L,COG1372L	replicative DNA helicase
61049	g-t	Rv0058	COG0305L,COG1372L	replicative DNA helicase
62174	a-c	Rv0058	COG0305L,COG1372L	replicative DNA helicase
62295	g-n	Rv0058	COG0305L,COG1372L	replicative DNA helicase
62301	g-n	Rv0058	COG0305L,COG1372L	replicative DNA helicase
62906	c-a	Rv0058	COG0305L,COG1372L	replicative DNA helicase
4400070	g-c	Rv3910	COG0515RTKL,COG0728R	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN

D - Cell division and chromosome partitioning

3864994	t-c	Rv3447c	COG1674D	PROBABLE CONSERVED MEMBRANE PROTEIN
3867395	t-a	Rv3447c	COG1674D	PROBABLE CONSERVED MEMBRANE PROTEIN
4379679	c-g	Rv3894c	COG1674D	POSSIBLE CONSERVED MEMBRANE PROTEIN

O - Posttranslational modification, protein turnover, chaperones

151284	c-t	Rv0125	COG0265O	PROBABLE SERINE PROTEASE PEPA (SERINE PROTEINASE) (MTB32A)
--------	-----	--------	----------	--

2376079	a-n	Rv2115c	COG0464O	Probable ATPase
4038286	g-a	Rv3596c	COG0542O	PROBABLE ATP-DEPENDENT PROTEASE ATP-BINDING SUBUNIT CLPC1
M - Cell envelope biogenesis				
3796788	c-t	Rv3382c	COG0761IM	PROBABLE LYTB-RELATED PROTEIN LYTB1
3856223	g-a	Rv3436c	COG0449M	D-fructose-6-phosphate amidotransferase
P - Inorganic ion transport and metabolism				
127014	c-a	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
127134	a-g	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
127290	g-a	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
128415	c-t	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
129147	g-c	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
2155167	c-g	Rv1908c	COG0376P	CATALASE-PEROXIDASE-PEROXYNITRITASE T KATG
T - Signal Transduction mechanisms				
3776705	c-t	Rv3365c	COG0642T	hypothetical protein
C - Energy production and conversion				
72735	g-a	Rv0066c	COG2838C	PROBABLE ISOCITRATE DEHYDROGENASE [NADP] ICD2 (OXALOSUCCINATE DECARBOXYLASE) (IDH) (NADP+-SPECIFIC ICDH) (IDP)
G - Carbohydrate transport and metabolism				
152409	g-a	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
153141	c-a	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
155087	c-g	Rv0127	COG3281G	hypothetical protein
E - Amino acid transport and metabolism				
76814	c-g	Rv0069c	COG1760E	PROBABLE L-SERINE DEHYDRATASE SDAA (L-SERINE DEAMINASE) (SDH) (L-SD)
84162	g-a	Rv0075	COG1168E	PROBABLE AMINOTRANSFERASE
138059	g-a	Rv0114	COG0241E	POSSIBLE D-ALPHA,BETA-D-HEPTOSE-1,7-BIPHOSPHATE PHOSPHATASE GMHB (D-GLYCERO-D-MANNO-HEPTOSE 7-PHOSPHATE KINASE)
1199546	g-a	Rv1075c	COG2755E	CONSERVED EXPORTED PROTEIN
4153387	t-g	Rv3709c	COG0527E	aspartate kinase
F - Nucleotide transport and mechnism				
3829769	t-c	Rv3410c	COG0516F	inositol-5-monophosphate dehydrogenase
H - Coenzyme metabolism				
142523	c-a	Rv0118c	COG0028EH	putative oxalyl-CoA decarboxylase
1777212	c-g	Rv1569	COG0156H	8-amino-7-oxononanoate synthase
2193247	g-a	Rv1940	COG0108H,COG0807H	Probable Riboflavin biosynthesis protein ribA1 (GTP cyclohydrolase II)
3714756	a-c	Rv3329	COG0161H	hypothetical protein
3798094	a-c	Rv3383c	COG0142H	POSSIBLE POLYPRENYL SYNTHETASE IDSB (POLYPRENYL TRANSFERASE) (POLYPRENYL DIPHOSPHATE SYNTHASE)

I - Lipid metabolism

3191026	g-a	Rv2881c	COG0575I	PROBABLE INTEGRAL MEMBRANE PHOSPHATIDATE CYTIDYLYLTRANSFERASE CDSA (CDP-DIGLYCERIDE SYNTHETASE) (CDP-DIGLYCERIDE PYROPHOSPHORYLASE) (CDP-DIACYLGLYCEROL SYNTHASE) (CDS) (CTP: PHOSPHATIDATE CYTIDYLYLTRANSFERASE) (CDP-DAG SYNTHASE) (CDP-DG SYNTHETASE)
3952799	g-a	Rv3516	COG1024I	enoyl-CoA hydratase
4011949	g-a	Rv3570c	COG1960I	POSSIBLE OXIDOREDUCTASE

Q - Secondary metabolites biosynthesis, transport and catabolism

37816	g-a	Rv0035	COG0318IQ	PROBABLE FATTY-ACID-CoA LIGASE FADD34 (FATTY-ACID-CoA SYNTHETASE) (FATTY-ACID-CoA SYNTHASE)
163413	c-g	Rv0136	COG2124Q	PROBABLE CYTOCHROME P450 138 CYP138
163436	g-t	Rv0136	COG2124Q	PROBABLE CYTOCHROME P450 138 CYP138
1001603	g-a	Rv0897c	COG1233Q	PROBABLE OXIDOREDUCTASE
2288846	c-g	Rv2043c	COG1335Q	PYRAZINAMIDASE/NICOTINAMIDAS PNCA (PZase)
2296041	g-c	Rv2048c	COG3321Q	Probable polyketide synthase pks12
2674577	g-c	Rv2383c	COG1020Q,COG3208Q	PHENYLOXAZOLINE SYNTHASE MBTB (PHENYLOXAZOLINE SYNTHETASE)
3817116	c-a	Rv3399	COG3315Q	hypothetical protein
3825559	c-g	Rv3406	COG2175Q	PROBABLE DIOXYGENASE
4293071	g-a	Rv3824c	COG1020Q	PROBABLE CONSERVED POLYKETIDE SYNTHASE ASSOCIATED PROTEIN PAPA1

R - General function prediction only

39949	g-a	Rv0037c	COG0477GEPR	PROBABLE CONSERVED INTEGRAL MEMBRANE PROTEIN
133533	t-c	Rv0110	COG0705R	PROBABLE CONSERVED INTEGRAL MEMBRANE PROTEIN
133673	g-c	Rv0110	COG0705R	PROBABLE CONSERVED INTEGRAL MEMBRANE PROTEIN
161991	t-c	Rv0134	COG0596R	POSSIBLE EPOXIDE HYDROLASE EPHF (EPOXIDE HYDRATASE) (ARENE-OXIDE HYDRATASE)
1944401	t-c	Rv1716	COG1878R	hypothetical protein
3718356	c-t	Rv3331	COG0477GEPR	PROBABLE SUGAR-TRANSPORT INTEGRAL MEMBRANE PROTEIN SUGI
3838870	a-g	Rv3420c	COG0456R	PROBABLE RIBOSOMAL-PROTEIN-ALANINE ACETYLTRANSFERASE RIMI (ACETYLATED ENZYME FOR N-TERMINAL OF RIBOSOMAL PROTEIN S18)

S - Function unknown

1995470	g-a	Rv1762c	COG0393S	hypothetical protein
2958387	c-n	Rv2631	COG1690S	hypothetical protein
3823158	a-t	Rv3403c	COG4529S	hypothetical protein

KZN4207 (DS) unique SNVs

Position	SNP	Gene	COG	Gene Description
----------	-----	------	-----	------------------

J - Translation, ribosomal structure and biogenesis

1917971	a-g	Rv1694	COG1189J	CYTOTOXIN HAEMOLYSIN HOMOLOGUE TLYA
---------	-----	--------	----------	-------------------------------------

K - Transcription

791485	g-a	Rv0691c	COG1309K	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN
3903529	g-c	Rv3484	COG1316K	POSSIBLE CONSERVED PROTEIN CPSA

L - DNA replication, recombination and repair

9039	a-g	Rv0006	COG0188L	DNA gyrase subunit A
9072	c-g	Rv0006	COG0188L	DNA gyrase subunit A
9295	t-g	Rv0006	COG0188L	DNA gyrase subunit A

D - Cell division and chromosome partitioning

3177967	g-a	Rv2866	COG2026JD	hypothetical protein
4348126	c-t	Rv3870	COG1674D	POSSIBLE CONSERVED TRANSMEMBRANE PROTEIN

O - Posttranslational modification, protein turnover, chaperones

235007	g-t	Rv0198c	COG3590O	PROBABLE ZINC METALLOPROTEASE
235040	c-g	Rv0198c	COG3590O	PROBABLE ZINC METALLOPROTEASE
235264	c-g	Rv0198c	COG3590O	PROBABLE ZINC METALLOPROTEASE
236210	t-g	Rv0198c	COG3590O	PROBABLE ZINC METALLOPROTEASE
302557	t-c	Rv0251c	COG0071O	HEAT SHOCK PROTEIN HSP (HEAT-STRESS-INDUCED RIBOSOME-BINDING PROTEIN A)

M - Cell envelope biogenesis

4233298	g-a	Rv3786c	COG0463M,COG0739M	hypothetical protein
4407903	g-a	Rv3919c	COG0357M	glucose-inhibited division protein B
4408155	a-c	Rv3919c	COG0357M	glucose-inhibited division protein B

P - Inorganic ion transport and metabolism

122108	a-g	Rv0103c	COG2217P	PROBABLE CATION-TRANSPORTER P-TYPE ATPASE B CTPB
127179	c-g	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
127212	t-c	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
127436	c-g	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
129114	g-t	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI

C - Energy production and conversion

73354	c-a	Rv0066c	COG2838C	PROBABLE ISOCITRATE DEHYDROGENASE [NADP] ICD2 (OXALOSUCCINATE DECARBOXYLASE) (IDH) (NADP+-SPECIFIC ICDH) (IDP)
73387	g-c	Rv0066c	COG2838C	PROBABLE ISOCITRATE DEHYDROGENASE [NADP] ICD2 (OXALOSUCCINATE DECARBOXYLASE) (IDH) (NADP+-SPECIFIC ICDH) (IDP)
73611	g-a	Rv0066c	COG2838C	PROBABLE ISOCITRATE DEHYDROGENASE [NADP] ICD2 (OXALOSUCCINATE DECARBOXYLASE) (IDH) (NADP+-SPECIFIC ICDH) (IDP)
90143	a-g	Rv0082	COG3260C	PROBABLE OXIDOREDUCTASE
279762	t-c	Rv0234c	COG1012C	succinic semialdehyde dehydrogenase
280494	c-a	Rv0234c	COG1012C	succinic semialdehyde dehydrogenase
4315383	t-c	Rv3842c	COG0584C	PROBABLE GLYCEROPHOSPHORYL DIESTER PHOSPHODIESTERASE GLPQ1 (GLYCEROPHOSPHODIESTER PHOSPHODIESTERASE)

G - Carbohydrate transport and metabolism

153035	a-g	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
154787	g-a	Rv0127	COG3281G	hypothetical protein
154820	a-c	Rv0127	COG3281G	hypothetical protein
155044	g-c	Rv0127	COG3281G	hypothetical protein
2812702	c-t	Rv2498c	COG2301G	PROBABLE CITRATE (PRO-3S)-LYASE (BETA SUBUNIT) CITE (CITRASE) (CITRATASE) (CITRITASE) (CITRIDESMOLASE) (CITRASE ALDOLASE)

F - Nucleotide transport and mechanism

278816	t-g	Rv0233	COG0208F	ribonucleotide-diphosphate reductase subunit beta
--------	-----	--------	----------	---

H - Coenzyme metabolism

307409	t-c	Rv0255c	COG1492H	cobyrinic acid synthase
1591661	c-t	Rv1415	COG0108H,COG0807H	bifunctional 3,4-dihydroxy-2-butanone 4-phosphate synthase/GTP cyclohydrolase II protein

I - Lipid metabolism

293335	g-t	Rv0243	COG0183I	acetyl-CoA acetyltransferase
294899	c-g	Rv0244c	COG1960I	PROBABLE ACYL-CoA DEHYDROGENASE FADE5
294932	t-a	Rv0244c	COG1960I	PROBABLE ACYL-CoA DEHYDROGENASE FADE5
295156	t-a	Rv0244c	COG1960I	PROBABLE ACYL-CoA DEHYDROGENASE FADE5
4255921	a-g	Rv3799c	COG4799I	PROBABLE PROPIONYL-CoA CARBOXYLASE BETA CHAIN 4 ACCD4 (PCCASE) (PROPANOYL-COA: CARBON DIOXIDE LIGASE)

Q - Secondary metabolites biosynthesis, transport and catabolism

37817	t-a	Rv0035	COG0318IQ	PROBABLE FATTY-ACID-CoA LIGASE FADD34 (FATTY-ACID-CoA SYNTHETASE) (FATTY-ACID-CoA SYNTHASE)
38041	g-a	Rv0035	COG0318IQ	PROBABLE FATTY-ACID-CoA LIGASE FADD34 (FATTY-ACID-CoA SYNTHETASE) (FATTY-ACID-CoA SYNTHASE)
82998	g-t	Rv0074	COG1228Q	hypothetical protein
171900	c-a	Rv0145	COG3315Q	hypothetical protein
171933	t-g	Rv0145	COG3315Q	hypothetical protein
172157	a-g	Rv0145	COG3315Q	hypothetical protein
173103	g-t	Rv0146	COG3315Q	hypothetical protein
2300545	a-t	Rv2048c	COG3321Q	Probable polyketide synthase pks12
2674094	g-a	Rv2383c	COG1020Q,COG3208Q	PHENYLOXAZOLINE SYNTHASE MBTB (PHENYLOXAZOLINE SYNTHETASE)
3973953	c-t	Rv3535c	COG4569Q	acetaldehyde dehydrogenase

R - General function prediction only

156722	c-a	Rv0129c	COG0627R	SECRETED ANTIGEN 85-C FBPC (85C) (ANTIGEN 85 COMPLEX C) (AG58C) (MYCOLYL TRANSFERASE 85C) (FIBRONECTIN-BINDING PROTEIN C)
291069	c-t	Rv0242c	COG1028IQR	3-ketoacyl-(acyl-carrier-protein) reductase
3938659	a-g	Rv3510c	COG2159R	hypothetical protein
4382053	t-c	Rv3896c	COG3953R	hypothetical protein

S - Function unknown

155990	g-a	Rv0128	COG3619S	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN
--------	-----	--------	----------	--

223941	t-c	Rv0192	COG1376S	hypothetical protein
242665	g-a	Rv0204c	COG0392S	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN
247517	a-c	Rv0207c	COG1432S	hypothetical protein
4174563	t-c	Rv3727	COG3349S	POSSIBLE OXIDOREDUCTASE
4187816	a-g	Rv3737	COG2966S,COG3610S	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN

KZN605 (XDR) unique SNVs

Position	SNP	Gene	COG	Gene Description
J - Translation, ribosomal structure and biogenesis				
3367764	g-a	Rv3009c	COG0064J	aspartyl/glutamyl-tRNA amidotransferase subunit B
K - Transcription				
162768	c-a	Rv0135c	COG1309K	POSSIBLE TRANSCRIPTIONAL REGULATORY PROTEIN
761109	a-g	Rv0667	COG0085K	DNA-directed RNA polymerase subunit beta
761160	t-c	Rv0667	COG0085K	DNA-directed RNA polymerase subunit beta
763122	t-c	Rv0667	COG0085K	DNA-directed RNA polymerase subunit beta
1526818	c-a	Rv1358	COG2114T,COG2771K,COG3903R	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN
3630177	c-a	Rv3249c	COG1309K	POSSIBLE TRANSCRIPTIONAL REGULATORY PROTEIN (PROBABLY TETR-FAMILY)
L - DNA replication, recombination and repair				
4012	t-c	Rv0003	COG1195L	recombination protein F
7361	g-c	Rv0006	COG0188L	DNA gyrase subunit A
7569	c-t	Rv0006	COG0188L	DNA gyrase subunit A
7584	g-c	Rv0006	COG0188L	DNA gyrase subunit A
890548	g-a	Rv0797	COG3547L	IS1547 transposase
890717	a-c	Rv0797	COG3547L	IS1547 transposase
1170403	c-a	Rv1047	COG3328L	PROBABLE TRANSPOSASE
1341623	g-t	Rv1199c	COG3328L	POSSIBLE TRANSPOSASE
1342062	t-c	Rv1199c	COG3328L	POSSIBLE TRANSPOSASE
1342580	t-c	Rv1199c	COG3328L	POSSIBLE TRANSPOSASE
1579137	g-n	Rv1402	COG1198L	primosome assembly protein PriA
2361603	c-g	Rv2101	COG0553KL	PROBABLE HELICASE HELZ
2361622	a-c	Rv2101	COG0553KL	PROBABLE HELICASE HELZ
2362040	c-a	Rv2101	COG0553KL	PROBABLE HELICASE HELZ
2439408	c-t	Rv2177c	COG3547L	POSSIBLE TRANSPOSASE
2439518	g-a	Rv2177c	COG3547L	POSSIBLE TRANSPOSASE
2828821	g-t	Rv2512c	COG3328L	IS1081 transposase
2829260	t-c	Rv2512c	COG3328L	IS1081 transposase
2829778	t-c	Rv2512c	COG3328L	IS1081 transposase
3117999	a-g	Rv2812	COG2801L	PROBABLE TRANSPOSASE
3381640	g-t	Rv3023c	COG3328L	PROBABLE TRANSPOSASE
3579494	c-g	Rv3202c	COG0210L,COG2887L	POSSIBLE ATP-DEPENDENT DNA HELICASE
3579497	a-t	Rv3202c	COG0210L,COG2887L	POSSIBLE ATP-DEPENDENT DNA HELICASE
3801266	g-a	Rv3387	COG3039L	POSSIBLE TRANSPOSASE

4077162	c-a	Rv3638	COG1484L	transposase
4182510	c-g	Rv3731	COG1793L	ATP-dependent DNA ligase
4182733	a-g	Rv3731	COG1793L	ATP-dependent DNA ligase
O - Posttranslational modification, protein turnover,chaperones				
151420	a-c	Rv0125	COG0265O	PROBABLE SERINE PROTEASE PEPA (SERINE PROTEINASE) (MTB32A)
M - Cell envelope biogenesis				
1361109	c-n	Rv1217c	COG3559M	PROBABLE TETRONASIN-TRANSPORT INTEGRAL MEMBRANE PROTEIN ABC TRANSPORTER
2939372	g-c	Rv2611c	COG1560M	lipid A biosynthesis lauroyl acyltransferase
3175701	t-c	Rv2864c	COG0768M	POSSIBLE PENICILLIN-BINDING LIPOPROTEIN
4408110	a-c	Rv3919c	COG0357M	glucose-inhibited division protein B
P - Inorganic ion transport and metabolism				
94151	a-g	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94274	g-c	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94316	t-g	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94372	c-t	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94395	g-a	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94414	a-g	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94422	g-a	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94428	a-g	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94464	t-c	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94477	g-t	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94484	t-c	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94496	c-g	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94557	c-a	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94562	c-g	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94621	g-t	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
94643	a-c	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
95123	t-a	Rv0086	COG0651CP	POSSIBLE HYDROGENASE HYCQ
99164	g-t	Rv0090	COG0601EP,COG3305S	POSSIBLE MEMBRANE PROTEIN
100708	c-g	Rv0092	COG2217P	PROBABLE CATION TRANSPORTER P-TYPE ATPASE A CTPA
129149	c-g	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
129579	t-g	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
129750	t-g	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
129829	c-a	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
129834	c-a	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
129848	c-g	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
129862	g-c	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
129892	g-c	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
129902	g-a	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
129911	g-c	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
129922	c-a	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
129991	t-c	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI

130077	t-a	Rv0107c	COG0474P	PROBABLE CATION-TRANSPORTER ATPASE I CTPI
756756	c-t	Rv0663	COG3119P	POSSIBLE ARYLSULFATASE ATSD (ARYL-SULFATE SULPHOHYDROLASE) (ARYLSULPHATASE)
3483698	g-a	Rv3116	COG0476H,COG0607P	PROBABLE MOLYBDENUM COFACTOR BIOSYNTHESIS PROTEIN MOEB2 (MPT-SYNTHASE SULFURYLASE) (MOLYBDOPTERIN SYNTHASE SULPHURYLASE)

C - Energy production and conversion

93715	t-g	Rv0085	COG4237C	POSSIBLE HYDROGENASE HYCP
93718	t-g	Rv0085	COG4237C	POSSIBLE HYDROGENASE HYCP
1292101	a-g	Rv1162	COG1140C	PROBABLE RESPIRATORY NITRATE REDUCTASE (BETA CHAIN) NARH

T - Singal Transduction mechanisms

24124	g-a	Rv0020c	COG1716T	hypothetical protein
3596543	c-t	Rv3220c	COG3920T	PROBABLE TWO COMPONENT SENSOR KINASE

G - Carbohydrate transport and metabolism

152673	c-t	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
152839	t-g	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
152895	t-c	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
152918	g-t	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
152923	c-a	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
152931	c-t	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
152937	c-g	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
152945	a-t	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
152951	a-c	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
152987	a-g	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
152991	c-g	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
153000	t-g	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
153007	g-t	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
153027	t-g	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
153080	a-c	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
153085	c-a	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
153144	c-g	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
153166	a-g	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
153600	g-t	Rv0126	COG0366G	TREHALOSE SYNTHASE TRES
154866	c-t	Rv0127	COG3281G	hypothetical protein
1630147	a-c	Rv1449c	COG0021G	transketolase
2470148	t-c	Rv2205c	COG1929G	hypothetical protein

E - Amino acid transport and metabolism

2246031	t-c	Rv2000	COG0493ER,COG0665E	hypothetical protein
---------	-----	--------	--------------------	----------------------

F - Nucleotide transport and mechnism

99781	c-t	Rv0091	COG0775F	PROBABLE BIFUNCTIONAL MTA/SAH NUCLEOSIDASE MTN: 5'-METHYLTHIOADENOSINE NUCLEOSIDASE (METHYLTHIOADENOSINE METHYLTHIORIBOHYDROLASE) + S-ADENOSYLHOMOCYSTEINE NUCLEOSIDASE (S-ADENOSYL-L-HOMOCYSTEINE HOMOCYSTEINYLRIBOHYDROLASE)
-------	-----	--------	----------	--

H - Coenzyme metabolism

2509139	g-c	Rv2236c	COG1270H	cobalamin biosynthesis protein
---------	-----	---------	----------	--------------------------------

I - Lipid metabolism

133976	c-a	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134032	g-c	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134039	a-c	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134042	g-t	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134055	t-c	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134060	g-a	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134064	t-a	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134074	a-c	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134084	c-g	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134088	t-a	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134124	g-c	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134137	c-a	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134143	c-t	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134148	t-g	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134217	a-c	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134222	a-g	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134281	c-t	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
134737	g-c	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
135215	c-g	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
136003	t-g	Rv0111	COG1835I	POSSIBLE TRANSMEMBRANE ACYLTRANSFERASE
159231	t-c	Rv0131c	COG1960I	PROBABLE ACYL-CoA DEHYDROGENASE FADE1
294011	t-g	Rv0244c	COG1960I	PROBABLE ACYL-CoA DEHYDROGENASE FADE5
2786951	a-g	Rv2482c	COG2937I	glycerol-3-phosphate acyltransferase
3606367	t-g	Rv3229c	COG3239I	POSSIBLE LINOLEOYL-CoA DESATURASE (DELTA(6)-DESATURASE)
3953252	t-c	Rv3516	COG1024I	enoyl-CoA hydratase

Q - Secondary metabolites biosynthesis, transport and catabolism

38710	g-c	Rv0035	COG0318IQ	PROBABLE FATTY-ACID-CoA LIGASE FADD34 (FATTY-ACID-CoA SYNTHETASE) (FATTY-ACID-CoA SYNTHASE)
922003	c-t	Rv0830	COG3315Q	hypothetical protein
1327889	g-a	Rv1186c	COG2508TQ	hypothetical protein
2300236	a-g	Rv2048c	COG3321Q	Probable polyketide synthase pks12
2300665	g-a	Rv2048c	COG3321Q	Probable polyketide synthase pks12
2300668	g-c	Rv2048c	COG3321Q	Probable polyketide synthase pks12
2300689	g-a	Rv2048c	COG3321Q	Probable polyketide synthase pks12

2300698	g-a	Rv2048c	COG3321Q	Probable polyketide synthase pks12
2302032	g-a	Rv2048c	COG3321Q	Probable polyketide synthase pks12
2674131	-	Rv2383c	COG1020Q,COG3208Q	PHENYLOXAZOLINE SYNTHASE MBTB (PHENYLOXAZOLINE SYNTHETASE)
2676709	-	Rv2384	COG1021Q	BIFUNCTIONAL ENZYME MBTA: SALICYL-AMP LIGASE (SAL-AMP LIGASE) + SALICYL-S-ArCP SYNTHETASE
4256911	t-c	Rv3800c	COG3319Q,COG3321Q	POLYKETIDE SYNTHASE PKS13
R - General function prediction only				
133375	t-g	Rv0110	COG0705R	PROBABLE CONSERVED INTEGRAL MEMBRANE PROTEIN
138824	c-g	Rv0115	COG2605R	POSSIBLE D-ALPHA-D-HEPTOSE-7-PHOSPHATE KINASE HDDA
1586248	c-t	Rv1410c	COG0477GEPR	AMINOGLYCOSIDES/TETRACYCLINE-TRANSPORT INTEGRAL MEMBRANE PROTEIN
2269779	t-c	Rv2024c	COG4889R	hypothetical protein
2808476	c-t	Rv2494	COG1848R	hypothetical protein
4306422	c-g	Rv3832c	COG0500QR	hypothetical protein
4306591	g-c	Rv3832c	COG0500QR	hypothetical protein
4306601	c-g	Rv3832c	COG0500QR	hypothetical protein
S - Function unknown				
14250	g-a	Rv0012	COG3879S	PROBABLE CONSERVED MEMBRANE PROTEIN
140368	t-g	Rv0116c	COG1376S	POSSIBLE CONSERVED MEMBRANE PROTEIN
1093405	a-g	Rv0978c	COG3391S	PE-PGRS FAMILY PROTEIN
1095230	g-n	Rv0980c	COG3391S	PE-PGRS FAMILY PROTEIN
1095354	g-n	Rv0980c	COG3391S	PE-PGRS FAMILY PROTEIN
1095428	g-n	Rv0980c	COG3391S	PE-PGRS FAMILY PROTEIN
1095479	g-n	Rv0980c	COG3391S	PE-PGRS FAMILY PROTEIN
1096397	c-t	Rv0980c	COG3391S	PE-PGRS FAMILY PROTEIN
3135911	g-c	Rv2828c	COG4293S	hypothetical protein

Supplementary table S2. Common single-nucleotide variations (SNVs) shared among the three KZN strains in comparison with H37Rv strain of *M. tuberculosis*

1435-4207-605 COMMON SNVs

Position	SNP	Gene	COG	Gene Description
J - Translation, ribosomal structure and biogenesis				
88485	g-c	Rv0079	COG1544J	hypothetical protein
146086	t-c	Rv0120c	COG0480J	elongation factor G
979703	g-c	Rv0881	COG0566J	POSSIBLE RRNA METHYLTRANSFERASE (RRNA METHYLASE)
1126888	g-c	Rv1007c	COG0143J	methionyl-tRNA synthetase
1411209	t-g	Rv1263	COG0154J	amidase
1446922	t-g	Rv1292	COG0018J	arginyl-tRNA synthetase
1457143	c-t	Rv1300	COG2890J	PROBABLE HEMK PROTEIN HOMOLOG HEMK
1512030	t-g	Rv1347c	COG1670J	hypothetical protein
1854299	t-c	Rv1644	COG0566J	Possible 23S rRNA methyltransferase tsnR

K - Transcription

46814	g-c	Rv0042c	COG1846K	POSSIBLE TRANSCRIPTIONAL REGULATORY PROTEIN (PROBABLY MARR-FAMILY)
194680	g-c	Rv0165c	COG1802K	POSSIBLE TRANSCRIPTIONAL REGULATORY PROTEIN (PROBABLY GNTR-FAMILY)
231113	c-g	Rv0195	COG2197TK	POSSIBLE TWO COMPONENT TRANSCRIPTIONAL REGULATORY PROTEIN (PROBABLY LUXR-FAMILY)
311612	g-t	Rv0260c	COG0745TK,COG1587H	bifunctional uroporphyrinogen-III synthetase/response regulator domain protein
563419	t-c	Rv0472c	COG1309K	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN (POSSIBLY TETR-FAMILY)
990000	g-c	Rv0890c	COG2771K,COG3903R	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN (PROBABLY LUXR-FAMILY)
991514	c-t	Rv0890c	COG2771K,COG3903R	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN (PROBABLY LUXR-FAMILY)
1142265	a-c	Rv1020	COG1197LK	PROBABLE TRANSCRIPTION-REPAIR COUPLING FACTOR MFD (TRCF)
1403031	t-c	Rv1255c	COG1309K	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN
1519846	c-g	Rv1353c	COG1309K	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN
1536250	g-t	Rv1364c	COG2172T,COG2202T,C OG2208TK	hypothetical protein
1734993	c-t	Rv1534	COG1309K	Probable transcriptional regulator
1899350	t-c	Rv1674c	COG0607P,COG0640K	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN
2007544	t-g	Rv1773c	COG1414K	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN
2264781	c-a	Rv2017	COG1396K,COG2856E	POSSIBLE TRANSCRIPTIONAL REGULATORY PROTEIN

L - DNA replication, recombination and repair

8039	g-a	Rv0006	COG0188L	DNA gyrase subunit A
9303	g-a	Rv0006	COG0188L	DNA gyrase subunit A
17607	g-c	Rv0015c	COG0515RTKL	TRANSMEMBRANE SERINE/THREONINE-PROTEIN KINASE A PKNA (PROTEIN KINASE A) (STPK A)
726702	c-a	Rv0631c	COG1330L	PROBABLE EXONUCLEASE V (GAMMA CHAIN) RECC (EXODEOXYRIBONUCLEASE V GAMMA CHAIN)(EXODEOXYRIBONUCLEASE V POLYPEPTIDE)
1047164	t-c	Rv0938	COG1793L,COG3285L	ATP-dependent DNA ligase
1127647	c-a	Rv1008	COG0084L	PROBABLE DEOXYRIBONUCLEASE TATD (YJJV PROTEIN)
1396921	t-c	Rv1251c	COG1112L,COG2251R	hypothetical protein
1428505	g-t	Rv1278	COG0419L,COG4717S	hypothetical protein
1740770	a-c	Rv1537	COG0389L	DNA polymerase IV
2347441	g-c	Rv2090	COG0258L	Probable 5'-3' exonuclease
2348445	c-g	Rv2090	COG0258L	Probable 5'-3' exonuclease
2621057	g-a	Rv2343c	COG0358L	DNA primase
2903462	c-t	Rv2578c	COG1533L	hypothetical protein
3238702	t-g	Rv2924c	COG0266L	formamidopyrimidine-DNA glycosylase

3425853	c-t	Rv3062	COG1793L	ATP-dependent DNA ligase
3426794	c-g	Rv3062	COG1793L	ATP-dependent DNA ligase
D - Cell division and chromosome partitioning				
2020562	a-t	Rv1783	COG1674D	PROBABLE CONSERVED MEMBRANE PROTEIN
2022867	t-c	Rv1784	COG1674D	hypothetical protein
3060353	g-c	Rv2748c	COG1674D	POSSIBLE CELL DIVISION TRANSMEMBRANE PROTEIN FTSK
3236229	c-a	Rv2922c	COG1196D	PROBABLE CHROMOSOME PARTITION PROTEIN SMC
O - Posttranslational modification, protein turnover,chaperones				
387352	g-a	Rv0319	COG2039O	pyrrolidone-carboxylate peptidase
420007	a-g	Rv0350	COG0443O	molecular chaperone DnaK
1100233	t-c	Rv0983	COG0265O	PROBABLE SERINE PROTEASE PEPD (SERINE PROTEINASE) (MTB32B)
1650071	a-g	Rv1462	COG0719O	hypothetical protein
2037715	t-c	Rv1798	COG0464O	hypothetical protein
2220511	t-g	Rv1977	COG0501O	hypothetical protein
3186859	t-g	Rv2874	COG0526OC,COG0785O	POSSIBLE INTEGRAL MEMBRANE C-TYPE CYTOCHROME BIOGENESIS PROTEIN DIPZ
M - Cell envelope biogenesis				
55552	c-t	Rv0050	COG0744M	PROBABLE BIFUNCTIONAL PENICILLIN-BINDING PROTEIN 1A/1B PONA1 (MUREIN POLYMERASE) (PBP1): PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE (PEPTIDOGLYCAN TGASE) + PENICILLIN-SENSITIVE TRANSPEPTIDASE (DD-TRANSPEPTIDASE)
535694	c-t	Rv0447c	COG2230M	PROBABLE CYCLOPROPANE-FATTY-ACYL-PHOSPHOLIPID SYNTHASE UFAA1 (CYCLOPROPANE FATTY ACID SYNTHASE) (CFA SYNTHASE)
796508	g-t	Rv0696	COG0463M	PROBABLE MEMBRANE SUGAR TRANSFERASE
1109974	a-g	Rv0993	COG1210M	PROBABLE UTP--GLUCOSE-1-PHOSPHATE URIDYLYLTRANSFERASE GALU (UDP-GLUCOSE PYROPHOSPHORYLASE) (UDPGP) (ALPHA-D-GLUCOSYL-1-PHOSPHATE URIDYLYLTRANSFERASE) (URIDINE DIPHOSPHOGLUCOSE PYROPHOSPHORYLASE)
1360208	t-c	Rv1217c	COG3559M	PROBABLE TETRONASIN-TRANSPORT INTEGRAL MEMBRANE PROTEIN ABC TRANSPORTER
1706118	t-c	Rv1514c	COG0463M	hypothetical protein
2415655	g-c	Rv2155c	COG0771M	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate synthetase
P - Inorganic ion transport and metabolism				
92198	t-g	Rv0083	COG0651CP	PROBABLE OXIDOREDUCTASE
511517	t-g	Rv0425c	COG0474P	POSSIBLE METAL CATION TRANSPORTING P-TYPE ATPASE CTPH
513256	t-c	Rv0425c	COG0474P	POSSIBLE METAL CATION TRANSPORTING P-TYPE ATPASE CTPH
637318	g-a	Rv0545c	COG0306P	PROBABLE LOW-AFFINITY INORGANIC PHOSPHATE TRANSPORTER INTEGRAL MEMBRANE PROTEIN PITA

1037011	t-c	Rv0930	COG0581P	PROBABLE PHOSPHATE-TRANSPORT INTEGRAL MEMBRANE ABC TRANSPORTER PSTA1
1040049	c-t	Rv0932c	COG0226P	PERIPLASMIC PHOSPHATE-BINDING LIPOPROTEIN PSTS2 (PBP-2) (PSTS2)
1375723	a-c	Rv1232c	COG2239P	hypothetical protein
1382627	t-c	Rv1239c	COG0598P	POSSIBLE MAGNESIUM AND COBALT TRANSPORT TRANSMEMBRANE PROTEIN CORA
1440468	c-g	Rv1286	COG0529P,COG2895P	bifunctional sulfate adenylyltransferase subunit 1/ adenylylsulfate kinase protein
1967236	c-a	Rv1739c	COG0659P	PROBABLE SULPHATE-TRANSPORT TRANSMEMBRANE PROTEIN ABC TRANSPORTER
2315668	g-a	Rv2059	COG0803P	hypothetical protein
2598399	a-g	Rv2326c	COG1122P	POSSIBLE TRANSMEMBRANE ATP-BINDING PROTEIN ABC TRANSPORTER
2695377	c-g	Rv2398c	COG4208P	PROBABLE SULFATE-TRANSPORT INTEGRAL MEMBRANE PROTEIN ABC TRANSPORTER CYSW
3440463	t-g	Rv3077	COG3119P	POSSIBLE HYDROLASE
3440467	g-c	Rv3077	COG3119P	POSSIBLE HYDROLASE
T - Singal Transduction mechanisms				
993345	a-c	Rv0891c	COG2114T	POSSIBLE TRANSCRIPTIONAL REGULATORY PROTEIN
1149550	c-t	Rv1028c	COG2205T	PROBABLE SENSOR PROTEIN KDPD
1150584	g-a	Rv1028c	COG2205T	PROBABLE SENSOR PROTEIN KDPD
1295417	g-a	Rv1165	COG1217T	POSSIBLE GTP-BINDING TRANSLATION ELONGATION FACTOR TYPA (TYROSINE PHOSPHORYLATED PROTEIN A) (GTP-BINDING PROTEIN)
1480023	g-t	Rv1318c	COG2114T	POSSIBLE ADENYLATE CYCLASE (ATP PYROPHOSPHATE-LYASE) (ADENYLYL CYCLASE)
1482626	t-c	Rv1320c	COG2114T	POSSIBLE ADENYLATE CYCLASE (ATP PYROPHOSPHATE-LYASE) (ADENYLYL CYCLASE)
1856776	g-c	Rv1647	COG2114T	hypothetical protein
3428916	c-a	Rv3063	COG1966T	PROBABLE CARBON STARVATION PROTEIN A HOMOLOG CSTA
3429201	t-g	Rv3063	COG1966T	PROBABLE CARBON STARVATION PROTEIN A HOMOLOG CSTA
C - Energy production and conversion				
74058	c-t	Rv0066c	COG2838C	PROBABLE ISOCITRATE DEHYDROGENASE [NADP] ICD2 (OXALOSUCCINATE DECARBOXYLASE) (IDH) (NADP+-SPECIFIC ICDH) (IDP)
403363	g-a	Rv0338c	COG0247C	PROBABLE IRON-SULFUR-BINDING REDUCTASE
403979	g-a	Rv0338c	COG0247C	PROBABLE IRON-SULFUR-BINDING REDUCTASE
404325	t-c	Rv0338c	COG0247C	PROBABLE IRON-SULFUR-BINDING REDUCTASE
558500	c-t	Rv0467	COG2224C	isocitrate lyase
589535	g-a	Rv0499	COG1155C	hypothetical protein
669397	t-c	Rv0575c	COG0654HC	hypothetical protein
861282	c-t	Rv0768	COG1012C	PROBABLE ALDEHYDE DEHYDROGENASE NAD DEPENDENT ALDA (ALDEHYDE DEHYDROGENASE [NAD+])

1074557	g-a	Rv0962c	COG1902C,COG4608E	POSSIBLE LIPOPROTEIN LPRP
1307597	c-g	Rv1175c	COG0446R,COG1902C	PROBABLE NADPH DEPENDENT 2,4-DIENOYL-COA REDUCTASE FADH (2,4-dienoyl coenzyme A reductase) (4-enoyl-CoA reductase)
1389737	g-a	Rv1248c	COG0508C,COG0567C	alpha-ketoglutarate decarboxylase
1759251	g-t	Rv1552	COG1053C	fumarate reductase flavoprotein subunit
1901815	a-g	Rv1677	COG05260C	PROBABLE CONSERVED LIPOPROTEIN DSBF
2055270	a-g	Rv1812c	COG1252C	PROBABLE DEHYDROGENASE
2422823	c-g	Rv2161c	COG2141C	hypothetical protein
2569402	c-a	Rv2298	COG0667C	hypothetical protein
3169583	c-t	Rv2858c	COG1012C	PROBABLE ALDEHYDE DEHYDROGENASE ALDC
3314628	c-t	Rv2962c	COG1819GC	POSSIBLE GLYCOSYL TRANSFERASE
G - Carbohydrate transport and metabolism				
680781	c-t	Rv0584	COG3537G	POSSIBLE CONSERVED EXPORTED PROTEIN
820482	g-t	Rv0727c	COG0235G	L-fuculose-phosphate aldolase
1144663	g-a	Rv1023	COG0148G	phosphopyruvate hydratase
1244699	t-c	Rv1121	COG0364G	glucose-6-phosphate 1-dehydrogenase
1613959	g-t	Rv1436	COG0057G	glyceraldehyde-3-phosphate dehydrogenase
1624790	c-g	Rv1446c	COG3429G	PUTATIVE OXPP CYCLE PROTEIN OPCA
1693560	a-g	Rv1502	COG2152G	hypothetical protein
1771357	c-t	Rv1564c	COG1523G	Probable Maltooligosyltrehalose synthase TreX
1804408	c-a	Rv1604	COG0483G	PROBABLE INOSITOL-MONOPHOSPHATASE IMPA (IMP)
2285250	c-a	Rv2039c	COG0395G	Probable sugar-transport integral membrane protein ABC transporter
2287120	a-g	Rv2041c	COG1653G	Probable sugar-binding lipoprotein
2734073	t-c	Rv2436	COG0524G	RIBOKINASE RBSK
E - Amino acid transport and metabolism				
501534	g-t	Rv0415	COG0665E	POSSIBLE THIAMINE BIOSYNTHESIS OXIDOREDUCTASE THIO
884552	a-g	Rv0790c	COG1305E	hypothetical protein
884673	c-t	Rv0790c	COG1305E	hypothetical protein
945213	g-a	Rv0848	COG0031E	POSSIBLE CYSTEINE SYNTHASE A CYSK2 (O-ACETYL SERINE SULFHYDRYLASE) (O-ACETYL SERINE (THIOL)-LYASE) (CSASE)
1057787	t-g	Rv0948c	COG1605E	hypothetical protein
1220679	t-c	Rv1093	COG0112E	serine hydroxymethyltransferase
1248977	t-c	Rv1125	COG0624E	hypothetical protein
1805947	c-t	Rv1606	COG0139E	phosphoribosyl-AMP cyclohydrolase
1901492	t-c	Rv1676	COG0626E	hypothetical protein
1931178	c-a	Rv1704c	COG1113E	PROBABLE D-SERINE/ALANINE/GLYCINE TRANSPORTER PROTEIN CYCA
2077252	g-a	Rv1832	COG0403E,COG1003E	glycine dehydrogenase
2128869	a-g	Rv1878	COG0174E	PROBABLE GLUTAMINE SYNTHETASE GLNA3 (GLUTAMINE SYNTHASE) (GS-I)
2128907	c-t	Rv1878	COG0174E	PROBABLE GLUTAMINE SYNTHETASE GLNA3 (GLUTAMINE SYNTHASE) (GS-I)

2245531	t-c	Rv2000	COG0493ER,COG0665E	hypothetical protein
2760151	a-g	Rv2458	COG2040E	homocysteine methyltransferase
2855258	a-g	Rv2531c	COG1982E	PROBABLE AMINO ACID DECARBOXYLASE
2889632	t-c	Rv2566	COG1305E,COG4196S	LONG CONSERVED HYPOTHETICAL PROTEIN
2911292	c-g	Rv2585c	COG0747E	POSSIBLE CONSERVED LIPOPROTEIN
3354895	c-t	Rv2996c	COG0111HE	PROBABLE D-3-PHOSPHOGLYCERATE DEHYDROGENASE SERA1 (PGDH)

F - Nucleotide transport and mechnism

458281	a-g	Rv0382c	COG0461F	orotate phosphoribosyltransferase
882256	t-c	Rv0787	COG0458EF	hypothetical protein
1068150	t-c	Rv0956	COG0299F	phosphoribosylglycinamide formyltransferase
3073867	t-c	Rv2764c	COG0207F	thymidylate synthase

H - Coenzyme metabolism

502588	c-g	Rv0417	COG2022H	thiazole synthase
1070701	t-c	Rv0958	COG1239H	POSSIBLE MAGNESIUM CHELATASE
1106421	t-c	Rv0989c	COG0142H	PROBABLE POLYPRENYL-DIPHOSPHATE SYNTHASE GRCC2 (POLYPRENYL PYROPHOSPHATE SYNTHETASE)
1797026	c-t	Rv1595	COG0029H	L-aspartate oxidase
2736433	c-a	Rv2438c	COG0171H,COG0388R	NAD synthetase

I - Lipid metabolism

327896	c-t	Rv0271c	COG1960I	PROBABLE ACYL-CoA DEHYDROGENASE FADE6
559051	a-c	Rv0468	COG1250I	3-hydroxybutyryl-CoA dehydrogenase
773808	g-a	Rv0673	COG1024I	enoyl-CoA hydratase
1083575	g-t	Rv0972c	COG1960I	PROBABLE ACYL-CoA DEHYDROGENASE FADE12
1264380	a-g	Rv1135A	COG0183I	POSSIBLE ACETYL-CoA ACETYLTRANSFERASE (ACETOACETYL-CoA THIOLASE)
1755518	g-t	Rv1551	COG2937I	glycerol-3-phosphate acyltransferase
1817975	a-t	Rv1618	COG1946I	Probable acyl-CoA thioesterase II tesB1
2116902	c-t	Rv1867	COG0183I	acetyl-CoA acetyltransferase
2642598	a-g	Rv2361c	COG0020I	LONG (C50) CHAIN Z-ISOPRENYL DIPHOSPHATE SYNTHASE (Z-DECAPRENYL DIPHOSPHATE SYNTHASE)
2818836	a-g	Rv2503c	COG2057I	PROBABLE SUCCINYL-COA:3-KETOACID-COENZYME A TRANSFERASE (BETA SUBUNIT) SCOB (3-OXO-ACID:COA TRANSFERASE) (OXCT B) (SUCCINYL CoA:3-OXOACID CoA-TRANSFERASE)
3165073	t-c	Rv2854	COG2267I	hypothetical protein

Q - Secondary metabolites biosynthesis, transport and catabolism

111359	t-c	Rv0101	COG1020Q,COG3320Q	PROBABLE PEPTIDE SYNTHETASE NRP (PEPTIDE SYNTHASE)
115999	t-g	Rv0101	COG1020Q,COG3320Q	PROBABLE PEPTIDE SYNTHETASE NRP (PEPTIDE SYNTHASE)
196521	c-t	Rv0166	COG0318IQ	acyl-CoA synthetase
203268	c-t	Rv0172	COG1463Q	MCE-FAMILY PROTEIN MCE1D
489934	g-c	Rv0405	COG3321Q	PROBABLE MEMBRANE BOUND POLYKETIDE SYNTHASE PKS6

685460	c-g	Rv0587	COG0767Q	CONSERVED HYPOTHETICAL INTEGRAL MEMBRANE PROTEIN YRBE2A
685607	t-c	Rv0587	COG0767Q	CONSERVED HYPOTHETICAL INTEGRAL MEMBRANE PROTEIN YRBE2A
686971	t-c	Rv0589	COG1463Q	MCE-FAMILY PROTEIN MCE2A
751946	t-c	Rv0655	COG1127Q	POSSIBLE RIBONUCLEOTIDE-TRANSPORT ATP-BINDING PROTEIN ABC TRANSPORTER MKL
752342	a-g	Rv0655	COG1127Q	POSSIBLE RIBONUCLEOTIDE-TRANSPORT ATP-BINDING PROTEIN ABC TRANSPORTER MKL
1315190	a-c	Rv1180	COG3321Q	PROBABLE POLYKETIDE BETA-KETOACYL SYNTHASE PKS3
1315883	g-a	Rv1181	COG3321Q	PROBABLE POLYKETIDE BETA-KETOACYL SYNTHASE PKS4
1327401	t-c	Rv1185c	COG0318IQ	acyl-CoA synthetase
1404168	t-g	Rv1256c	COG2124Q	PROBABLE CYTOCHROME P450 130 CYP130
1570565	c-a	Rv1394c	COG2124Q	PROBABLE CYTOCHROME P450 132 CYP132
1692140	a-c	Rv1501	COG5285Q	hypothetical protein
2210643	t-c	Rv1967	COG1463Q	MCE-FAMILY PROTEIN MCE3B
2216247	c-g	Rv1971	COG1463Q	MCE-FAMILY PROTEIN MCE3F
2518918	g-a	Rv2245	COG0304IQ	3-oxoacyl-(acyl carrier protein) synthase II
2660318	c-g	Rv2379c	COG1020Q	PEPTIDE SYNTHETASE MBTF (PEPTIDE SYNTHASE)
2821341	c-t	Rv2505c	COG0318IQ	acyl-CoA synthetase
3103681	t-c	Rv2794c	COG2977Q	hypothetical protein
3223619	a-g	Rv2915c	COG1228Q	hypothetical protein
3247315	c-g	Rv2931	COG3321Q	PHENOLPTHIOCEROL SYNTHESIS TYPE-I POLYKETIDE SYNTHASE PPSA
3261379	a-g	Rv2933	COG3321Q	PHENOLPTHIOCEROL SYNTHESIS TYPE-I POLYKETIDE SYNTHASE PPSC
3269580	a-g	Rv2935	COG1020Q,COG3321Q	PHENOLPTHIOCEROL SYNTHESIS TYPE-I POLYKETIDE SYNTHASE PPSE
3270783	a-g	Rv2935	COG1020Q,COG3321Q	PHENOLPTHIOCEROL SYNTHESIS TYPE-I POLYKETIDE SYNTHASE PPSE
3420055	g-a	Rv3059	COG2124Q	PROBABLE CYTOCHROME P450 136 CYP136
3456665	a-g	Rv3089	COG0318IQ	PROBABLE CHAIN-FATTY-ACID-CoA LIGASE FADD13 (FATTY-ACYL-CoA SYNTHETASE)
4072239	g-a	Rv3633	COG5285Q	hypothetical protein
R - General function prediction only				
147945	c-t	Rv0121c	COG3576R	hypothetical protein
157291	c-t	Rv0129c	COG0627R	SECRETED ANTIGEN 85-C FBPC (85C) (ANTIGEN 85 COMPLEX C) (AG58C) (MYCOLYL TRANSFERASE 85C) (FIBRONECTIN-BINDING PROTEIN C)
162580	g-a	Rv0134	COG0596R	POSSIBLE EPOXIDE HYDROLASE EPHF (EPOXIDE HYDRATASE) (ARENE-OXIDE HYDRATASE)
541200	a-g	Rv0450c	COG2409R	PROBABLE CONSERVED TRANSMEMBRANE TRANSPORT PROTEIN MMPL4
573261	a-g	Rv0484c	COG4221R	PROBABLE SHORT-CHAIN TYPE OXIDOREDUCTASE
597815	a-g	Rv0507	COG2409R	PROBABLE CONSERVED TRANSMEMBRANE TRANSPORT PROTEIN MMPL2

598474	g-a	Rv0507	COG2409R	PROBABLE CONSERVED TRANSMEMBRANE TRANSPORT PROTEIN MMPL2
775638	t-c	Rv0676c	COG2409R	PROBABLE CONSERVED TRANSMEMBRANE TRANSPORT PROTEIN MMPL5
880561	g-t	Rv0785	COG3573R	putative FAD-binding dehydrogenase
949534	t-c	Rv0853c	COG3961GHR	PROBABLE PYRUVATE OR INDOLE-3-PYRUVATE DECARBOXYLASE PDC
1025105	t-c	Rv0919	COG0454KR	hypothetical protein
1073032	t-g	Rv0959	COG4867R	hypothetical protein
1324723	c-a	Rv1184c	COG3008R	POSSIBLE EXPORTED PROTEIN
1325164	c-t	Rv1184c	COG3008R	POSSIBLE EXPORTED PROTEIN
1689348	c-t	Rv1498c	COG0500QR	PROBABLE METHYLTRANSFERASE
1762614	c-t	Rv1557	COG2409R	PROBABLE CONSERVED TRANSMEMBRANE TRANSPORT PROTEIN MMPL6
1798354	g-a	Rv1597	COG0500QR	hypothetical protein
1847918	c-g	Rv1639c	COG0627R	hypothetical protein
1848183	c-t	Rv1639c	COG0627R	hypothetical protein
1929251	t-g	Rv1703c	COG4122R	Probable catechol-o-methyltransferase
1943591	c-t	Rv1716	COG1878R	hypothetical protein
2143327	g-c	Rv1895	COG1063ER	POSSIBLE DEHYDROGENASE
2282786	c-t	Rv2037c	COG1752R	POSSIBLE CONSERVED TRANSMEMBRANE PROTEIN
2283029	a-g	Rv2037c	COG1752R	POSSIBLE CONSERVED TRANSMEMBRANE PROTEIN
2484254	t-g	Rv2216	COG1090R	hypothetical protein
2586126	a-g	Rv2314c	COG0312R	hypothetical protein
3029609	g-a	Rv2716	COG0384R	hypothetical protein
3041870	g-t	Rv2729c	COG0697GER	PROBABLE CONSERVED INTEGRAL MEMBRANE ALANINE VALINE AND LEUCINE RICH PROTEIN
3065823	g-a	Rv2752c	COG0595R	hypothetical protein
3108673	a-c	Rv2800	COG2936R	POSSIBLE HYDROLASE
3351925	t-c	Rv2994	COG0477GEPR	PROBABLE CONSERVED INTEGRAL MEMBRANE PROTEIN

S - Function unknown

14784	t-c	Rv0012	COG3879S	PROBABLE CONSERVED MEMBRANE PROTEIN
69988	g-a	Rv0064	COG1615S	hypothetical protein
70815	a-g	Rv0064	COG1615S	hypothetical protein
310972	g-a	Rv0259c	COG2138S	hypothetical protein
390827	t-c	Rv0323c	COG2120S	hypothetical protein
426908	a-c	Rv0355c	COG1357S	PPE FAMILY PROTEIN
444350	g-t	Rv0366c	COG4185S	hypothetical protein
551524	a-c	Rv0459	COG3564S	hypothetical protein
563576	a-g	Rv0473	COG2733S	POSSIBLE CONSERVED TRANSMEMBRANE PROTEIN
572467	g-a	Rv0483	COG1376S	PROBABLE CONSERVED LIPOPROTEIN LPRQ
1177993	t-c	Rv1056	COG2343S	hypothetical protein
1178115	t-c	Rv1056	COG2343S	hypothetical protein

1370736	t-c	Rv1227c	COG3402S	PROBABLE TRANSMEMBRANE PROTEIN
1815603	g-a	Rv1615	COG2314S	Probable hypothetical membrane protein
1967843	c-a	Rv1740	COG4423S	hypothetical protein
2053681	c-t	Rv1811	COG1285S	POSSIBLE Mg ²⁺ TRANSPORT P-TYPE ATPASE C MGTC
2071575	c-t	Rv1825	COG3879S	hypothetical protein
2499725	g-a	Rv2226	COG5607S	hypothetical protein
2503624	c-t	Rv2230c	COG0327S,COG3323S	hypothetical protein
2509721	a-g	Rv2237	COG3662S	hypothetical protein
2891266	c-t	Rv2567	COG2307S,COG2308S	CONSERVED HYPOTHETICAL ALANINE AND LEUCINE RICH PROTEIN
2891727	a-g	Rv2567	COG2307S,COG2308S	CONSERVED HYPOTHETICAL ALANINE AND LEUCINE RICH PROTEIN
2893237	c-a	Rv2568c	COG4307S	hypothetical protein
3133535	t-c	Rv2825c	COG4293S	hypothetical protein
3460985	g-a	Rv3092c	COG2354S	PROBABLE CONSERVED INTEGRAL MEMBRANE PROTEIN

Supplementary table S3. Common single-nucleotide variations (SNVs) shared between multidrug-resistant (MDR) and extensively drug-resistant (XDR) KZN strains in comparison with the H37Rv strain of *M. tuberculosis*

Common and unique SNV between MDR and XDR in comparison between <i>M. tuberculosis</i> H37Rv				
Position	SNP	Gene	COG	Gene Description
21794	g-a	Rv0018c	COG0631T	POSSIBLE SERINE/THREONINE PHOSPHATASE PPP
27462	c-g	intergenic	-	-----
177856	g-a	Rv0151c	-	PE FAMILY PROTEIN
206338	t-c	Rv0174	COG1463Q	MCE-FAMILY PROTEIN MCE1F
212352	c-t	Rv0181c	COG1741R	hypothetical protein
242298	c-g	Rv0204c	COG0392S	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN
265553	a-c	Rv0222	COG1024I	enoyl-CoA hydratase
278680	c-g	Rv0233	COG0208F	ribonucleotide-diphosphate reductase subunit beta
403919	c-t	Rv0338c	COG0247C	PROBABLE IRON-SULFUR-BINDING REDUCTASE
479631	g-t	Rv0400c	COG1960I	ACYL-CoA DEHYDROGENASE FADE7
498556	c-a	Rv0412c	-	POSSIBLE CONSERVED MEMBRANE PROTEIN
534690	c-t	Rv0446c	COG3752S	POSSIBLE CONSERVED TRANSMEMBRANE PROTEIN
610119	t-g	intergenic	-	-----
648001	t-g	Rv0556	-	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN
664248	c-t	Rv0571c	COG0596R,COG1926R,COG2267I	hypothetical protein
690464	t-g	Rv0591	COG1463Q	MCE-FAMILY PROTEIN MCE2C
754185	a-g	Rv0658c	COG1266R	PROBABLE CONSERVED INTEGRAL MEMBRANE PROTEIN
754197	c-t	Rv0658c	COG1266R	PROBABLE CONSERVED INTEGRAL MEMBRANE PROTEIN
812807	g-a	Rv0717	COG0199J	30S ribosomal protein S14
817855	c-g	Rv0724A	-	hypothetical protein

857695	a-g	Rv0764c	COG2124Q	CYTOCHROME P450 51 CYP51 (CYPL1) (P450-L1A1) (STEROL 14-ALPHA DEMETHYLASE) (LANOSTEROL 14-ALPHA DEMETHYLASE) (P450-14DM)
863804	t-c	Rv0770	COG2084I	PROBABLE DEHYDROGENASE/REDUCTASE
885541	g-c	Rv0791c	COG2141C	hypothetical protein
893732	t-g	Rv0800	COG1362E	putative aminopeptidase 2
903549	t-c	Rv0808	COG0034F	amidophosphoribosyltransferase
903912	t-c	Rv0809	COG0150F	phosphoribosylaminoimidazole synthetase
906856	a-g	Rv0812	COG0115EH	4-amino-4-deoxychorismate lyase
919392	g-c	Rv0825c	COG1309K	hypothetical protein
921812	c-g	Rv0829	COG0675L	POSSIBLE TRANSPOSASE (FRAGMENT)
1037910	c-t	Rv0930	COG0581P	PROBABLE PHOSPHATE-TRANSPORT INTEGRAL MEMBRANE ABC TRANSPORTER PSTA1
1068431	a-g	Rv0957	COG0138F	bifunctional phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase
1075278	t-c	intergenic	-	-----
1077311	a-g	Rv0966c	-	hypothetical protein
1103248	c-t	Rv0987	COG0577V	PROBABLE ADHESION COMPONENT TRANSPORT TRANSMEMBRANE PROTEIN ABC TRANSPORTER
1224366	t-c	intergenic	-	-----
1276587	c-g	Rv1148c	-	hypothetical protein
1331695	a-c	Rv1188	COG0506E	PROBABLE PROLINE DEHYDROGENASE
1332534	c-a	Rv1189	COG1595K	RNA polymerase sigma factor SigI
1374064	t-c	Rv1230c	COG2951M	POSSIBLE MEMBRANE PROTEIN
1393625	a-g	Rv1249c	-	POSSIBLE MEMBRANE PROTEIN
1480971	t-c	Rv1319c	COG2114T	POSSIBLE ADENYLATE CYCLASE (ATP PYROPHOSPHATE-LYASE) (ADENYLYL CYCLASE)
1481184	a-c	Rv1319c	COG2114T	POSSIBLE ADENYLATE CYCLASE (ATP PYROPHOSPHATE-LYASE) (ADENYLYL CYCLASE)
1573659	t-g	Rv1396c	-	PE-PGRS FAMILY PROTEIN
1651305	c-t	Rv1463	COG0396O	PROBABLE CONSERVED ATP-BINDING PROTEIN ABC TRANSPORTER
1673431	t-a	intergenic	-	-----
1718760	c-t	Rv1524	COG1819GC	Probable glycosyltransferase
1760291	a-g	Rv1554	COG3029C	PROBABLE FUMARATE REDUCTASE [MEMBRANE ANCHOR SUBUNIT] FRDC (FUMARATE DEHYDROGENASE) (FUMARIC HYDROGENASE)
1762253	g-a	Rv1557	COG2409R	PROBABLE CONSERVED TRANSMEMBRANE TRANSPORT PROTEIN MMPL6
1778429	t-c	Rv1570	COG0132H	dithiobiotin synthetase
1803264	g-a	Rv1602	COG0118E	imidazole glycerol phosphate synthase subunit HisH
1885771	g-a	Rv1662	COG3321Q	Probable polyketide synthase pks8
1895173	c-t	Rv1668c	COG0488R	PROBABLE FIRST PART OF MACROLIDE-TRANSPORT ATP-BINDING PROTEIN ABC TRANSPORTER
1933987	g-a	intergenic	-	-----
1952688	g-c	Rv1726	COG0277C	PROBABLE OXIDOREDUCTASE
1968639	t-a	Rv1742	-	hypothetical protein

2096185	a-g	Rv1846c	COG3682K	POSSIBLE TRANSCRIPTIONAL REGULATORY PROTEIN
2097989	a-c	Rv1850	COG0804E	urease subunit alpha
2134214	t-c	Rv1884c	-	PROBABLE RESUSCITATION-PROMOTING FACTOR RPFC
2221795	c-t	Rv1979c	COG0531E	POSSIBLE CONSERVED PERMEASE
2253849	g-a	Rv2006	COG0637R,COG1554G ,COG1877G	PROBABLE TREHALOSE-6-PHOSPHATE PHOSPHATASE OTSBI (TREHALOSE-PHOSPHATASE) (TPP)
2260524	c-t	intergenic	-	-----
2270101	a-g	Rv2024c	COG4889R	hypothetical protein
2401401	g-t	Rv2141c	COG0624E	hypothetical protein
2437970	g-a	Rv2176	COG0515RTKL	PROBABLE TRANSMEMBRANE SERINE/THREONINE- PROTEIN KINASE L PKNL (PROTEIN KINASE L) (STPK L)
2502072	c-t	Rv2228c	COG0328L,COG0406G	hypothetical protein
2505918	a-g	intergenic	-	-----
2531741	a-g	Rv2258c	COG0500QR	Possible transcriptional regulatory protein
2654370	g-a	Rv2374c	COG1420K	heat-inducible transcription repressor
2656224	a-g	Rv2377c	COG3251S	PUTATIVE CONSERVED PROTEIN MBTH
2665147	g-c	Rv2380c	COG1020Q	PEPTIDE SYNTHETASE MBTE (PEPTIDE SYNTHASE)
2830524	c-a	Rv2513	-	hypothetical protein
2865759	a-g	Rv2542	-	hypothetical protein
2865881	t-c	Rv2542	-	hypothetical protein
2888200	t-c	Rv2566	COG1305E,COG4196S	LONG CONSERVED HYPOTHETICAL PROTEIN
2894207	g-a	Rv2569c	COG1305E	hypothetical protein
2943410	t-c	Rv2614A	COG0149G	hypothetical protein
3080794	a-g	Rv2771c	COG0655R	hypothetical protein
3101118	g-t	Rv2791c	COG0675L	PROBABLE TRANSPOSASE
3108054	c-t	Rv2799	-	PROBABLE MEMBRANE PROTEIN
3205977	a-c	Rv2896c	COG0758LU	hypothetical protein
3218342	g-a	Rv2911	COG1686M	PROBABLE D-ALANYL-D-ALANINE CARBOXYPEPTIDASE DACB2 (PENICILLIN-BINDING PROTEIN) (DD-PEPTIDASE) (DD-CARBOXYPEPTIDASE) (PBP) (DD-TRANSPEPTIDASE) (SERINE-TYPE D- ALA-D-ALA CARBOXYPEPTIDASE) (D-AMINO ACID HYDROLASE)
3228142	g-t	Rv2917	COG1061KL	CONSERVED HYPOTHETICAL ALANINE AND ARGININE RICH PROTEIN
3256493	a-g	Rv2933	COG3321Q	PHENOLPTHIOLCEROL SYNTHESIS TYPE-I POLYKETIDE SYNTHASE PPSC
3272996	a-g	Rv2936	COG1131V	PROBABLE DAUNORUBICIN-DIM-TRANSPORT ATP- BINDING PROTEIN ABC TRANSPORTER DRRA
3338602	g-c	Rv2982c	COG0240C	NAD(P)H-dependent glycerol-3-phosphate dehydrogenase
3402815	c-t	Rv3042c	COG0560E,COG3830T	phosphoserine phosphatase
3454985	c-a	Rv3088	-	hypothetical protein
3462134	g-c	Rv3093c	COG2141C	HYPOTHETICAL OXIDOREDUCTASE