

Supplementary data:

Table 1. Oligonucleotide primers used for SNP verification by Sanger sequencing.

Oligo	5' – 3' sequence	Product size (bp)	Annealing temperature (°C)
<i>katG_L</i>	gccggtcaagaagaagtacg	591	67
<i>katG_R</i>	ctcttcgtcagctcccactc		
<i>rpoB_L^a</i>	ctgatcaaaaccagatccg	440	64
<i>rpoB_R^a</i>	tacacgatctcgtcctaac		
<i>embB_L^b</i>	cgacgccgtggtgatattcg	863	70
<i>embB_R^b</i>	ccacgctgggaattcgcttg		
<i>pncA_L^c</i>	ggcgtcatggaccctatc	670	64
<i>pncA_R^c</i>	caacagttcatcccggttc		
<i>rpoC_L</i>	agaacatcaagagcgaag	223	66
<i>rpoC_R</i>	atctggtcaccgtcgaagtc		
<i>sigA_L</i>	gcagatacgcactgaaa	947	66
<i>sigA_R</i>	ggctagctcgaccttctct		
<i>rpoB_L</i>	agctgagccaattcatggac	239	65
<i>rpoB_R</i>	cgtttcgatgaacccgaac		

^aAo W., Aldous S., Woodruff E., Hicke B., Rea L., Kreiswirth B. and Jenison R. 2012 Rapid detection of *rpoB* gene mutations conferring rifampin resistance in *Mycobacterium tuberculosis*. *J. Clin. Microbiol.* 50, 2433-2440.

^bBakuła Z., Napiórkowska A., Bielecki J. *et al.* 2013 Mutations in the *embB* gene and their association with ethambutol resistance in multidrug-resistant *Mycobacterium tuberculosis* clinical isolates from Poland. *Biomed. Res. Int.* (in press) <http://dx.doi.org/10.1155/2013/167954>.

^cAlexander D. C., Ma J. H., Guthrie J. L. *et al.* 2012 Gene sequencing for routine verification of pyrazinamide resistance in *Mycobacterium tuberculosis*: a role for *pncA* but not *rpsA*. *J. Clin. Microbiol.* 50, 3726-3728.

Table 2. Summary statistics clinical strains mapped to closely-related reference genomes.

	V9124 ^a	V1435 ^d	KZN605 ^d	X162 ^d	B910 ^b	R35 ^e	R490 ^c	R271 ^f
Mapped reads	1,830,877	937,477	999,617	1,055,628	899,721	2,018,291	1,406,634	2,681,225
% Of total reads	98.88	98.55	98.94	99.10	98.19	98.75	97.70	98.41
Average read length	265.15	232.95	225.41	235.45	230.76	243.35	188.51	193.25
Unmapped reads	20,759	13,837	10,689	9,542	16,555	25,461	33,112	43,371
% Of total reads	1.12	1.45	1.06	0.90	1.81	1.25	2.30	1.59
Average read length	285.21	268.05	278.23	282.43	261.55	274.64	198.23	207.32

^a Mapped to the KZN-V4207 reference genome (4,394,985 bp)

^b Mapped to the HN878 reference genome (4,404,672 bp)

^c Mapped to the F11 reference genome (4,424,435 bp)

Table 3. Summary statistics for clinical strains mapped to H37Rv reference genome (4,411,532 bp).

	V9124	V1435	KZN605	X162	B910	R35	R490	R271	R104	R443
Mapped reads	1,827,320	935,638	997,781	1,053,750	901,327	2,018,291	1,401,910	2,672,172	749,682	765,817
% Of total reads	98.69	98.35	98.76	98.93	98.37	98.75	97.37	98.08	98.53	98.24
Average read length	265.15	232.93	225.40	235.44	230.88	243.39	188.50	193.23	228.92	234.04
Unmapped reads	24,316	15,676	12,525	11,420	14,949	25,461	37,836	52,424	11,196	13,749
% Of total reads	1.31	1.65	1.24	1.07	1.63	1.25	2.63	1.92	1.47	1.76
Average read length	282.44	265.20	270.97	275.33	258.15	271.56	197.38	205.96	253.78	263.66

Table 4. Polymorphisms in re-sequenced V1435 relative to previously sequenced KZN 1435

Gene	Name	Function	Coding region change ^a	Amino acid change ^b
TBMG_00302		Predicted protein	1701_1702 ins G	P568fs
TBMG_00540	<i>menA</i>	1,4-Dihydroxy-2-naphthoate octaprenyltransferase	701_702 ins C	A234fs
TBMG_00782		Conserved hypothetical protein	547 del C	H183fs
TBMG_00937		Transferase	1139_1140 ins C	T380fs
TBMG_01038	<i>ppsD</i>	Phenolphthiocerol synthesis type-I polyketide synthase	3823 del A	I1275fs
TBMG_01735	<i>kasA</i>	3-Oxoacyl-[acyl-carrier protein] synthase 1	1001_1002 ins C	A334fs
TBMG_02215	<i>malQ</i>	4-Alpha-glucanotransferase	1342 del C	P448fs
TBMG_02275	<i>folP2</i>	Dihydropteroate synthase 2	632 del C	P211fs
TBMG_04073		Conserved hypothetical protein	809_810 ins G	A270fs
TBMG_04086	<i>trcS</i>	Two component system sensor histidine kinase	873_874 ins C	P292fs
TBMG_03038		Conserved hypothetical protein	26delG	R9fs
TBMG_03046		Monooxygenase	65_66 ins C	R22fs
TBMG_04096		Monooxygenase	1173 del C	P391fs
TBMG_03128		Conserved hypothetical protein	1378_1379 ins C	A460fs
TBMG_03778		Conserved hypothetical protein	468 del G	G156fs

^a del, deletion; ins, insertion^b fs, frameshift mutation

Table 5. Polymorphisms in re-sequenced KZN605 relative to previously sequenced KZN 605.

Gene	Function	Nucleotide change ^a	Amino acid change ^b
TBXG_000209	Conserved hypothetical protein	62delG	G21fs
TBXG_001385	Conserved hypothetical protein	863delT	F288fs
TBXG_001918	rifB protein	425delA	E1419fs
TBXG_002105	Conserved hypothetical protein	293delG	R98fs
TBXG_002696	Conserved hypothetical protein	552delA	K184fs
TBXG_003461	Conserved hypothetical protein	799C>T	A267T
TBXG_003898	Conserved hypothetical protein	445T>G	N149H

^a del, deletion^b fs, frameshift mutation