Supplemental Material

Allele (# repeats)	MS1	MS2	MS3	MS4	MS5	MS6	MS7	MS8	MS9
2					13				
3					12				
4	1				2	1			
5	7	2						4	
6		9						24	
7		31	1	3		1		3	
8	17	43	21	33	1	7		2	2
9	41	1	47	33	4	11	5	10	5
10	17			9	41	16	30	15	56
11	6			3	9	20	28	5	10
12				1	2	9	17	1	6
13			2	1		10	4	1	1
14			7			1			
15			4					3	
16			1					3	
N ✧	89	86	83	83	84	76	84	71	80
Alleles ✶	6	5	7	7	8	9	5	11	6
Range ✳	4-11	5-9	7-16	7-13	2-12	4-14	9-13	5-16	8-13
Mean H_e	0.688	0.586	0.667	0.607	0.620	0.842	0.709	0.789	0.449
% Amp †	97.8	94.5	91.2	91.2	92.3	83.5	92.3	78.0	87.9
$(Type)_N$	**TA**	**GAT**	**AT**	**TAT**	**AT**	**TA**	**TA**	**TA**	**TA**
Domain ‡	**3'UTR**	**Coding**	**5'UTR**	**3'UTR**	**N/A**	**Intron**	**Intron**	**3'UTR**	**3'UTR**
Gene Name	**Undetermined**	**RING-finger membrane protein**	**RNA-binding protein ISN1 (putative)**	**Sec7-domain containing protein**	**Intergenic**	**Aspartate aminotransferase**	**Serine hydroxymethyl transferase**	**E3 ubiquitin protein ligase (putative)**	**ATPase WRNIP1 homolog**

Table S1. Summary characteristic by locus: number of isolates typed, range and frequency of alleles, expected heterozygosity. ✧Total number of *Pneumocystis jirovecii* isolates. ✶Number of unique alleles. ✳Range of repeat lengths detected. †A measure of locus genotyping robustness: the per-locus success rate for PCR and capillary electrophoresis. ‡Microsatellite location within predicted gene.

Population Type	Sample ID	*dhps* Genotype*	MS9 Genotype†
Uganda Single Mutant	UG02	THR-ARG-SER	10
	UG03	THR-ARG-SER	10
	UG04	THR-ARG-SER	10
	UG06	THR-ARG-SER	10
	UG07	THR-ARG-SER	10
	UG08	THR-ARG-SER	9
	UG09	THR-ARG-SER	9
	UG12	THR-ARG-SER	10
	UG13	THR-ARG-SER	10
Uganda Double Mutant	UG10	ALA-ARG-SER	10
	UG11	ALA-ARG-SER	10
San Francisco Single Mutant	SF13	ALA-ARG-PRO	11
	SF47-1	THR-ARG-SER	12
	SF47-2	THR-ARG-SER	11
San Francisco – Double Mutant	SF05	ALA-ARG-SER	10
	SF11	ALA-ARG-SER	10
	SF12	ALA-ARG-SER	10
	SF19	ALA-ARG-SER	10
	SF22	ALA-ARG-SER	10
	SF27	ALA-ARG-SER	10
	SF30	ALA-ARG-SER	12
	SF48-1	ALA-ARG-SER	9
	SF48-2	ALA-ARG-SER	9
	SF49-1	ALA-ARG-SER	10
	SF49-2	ALA-ARG-SER	8

Table S2. Microsatellite #9 (MS9)-*dihydropteroate synthase* (*dhps*) haplotypes among Ugandan and San Franciscan isolates. 11 samples from Uganda and 14 samples from San Francisco had unambiguous and non-mixed genotype calls at both MS9 and *dhps*, allowing the construction of two-locus haplotypes. *Wildtype amino-acid sequence at the *dhps* locus is THR-ARG-PRO. †Genotype calls represent absolute number of dinucleotide repeats.

Simpson's D from H_e	**0.793**	**0.972**	**0.991**	**0.997**	**0.999**	**>0.999**	**>0.999**	**>0.999**
	MS6	MS6	MS6	MS6	MS6	MS6	MS6	MS6
		MS8	MS8	MS8	MS8	MS8	MS8	MS8
			MS7	MS7	MS7	MS7	MS7	MS7
				MS1	MS1	MS1	MS1	MS1
					MS3	MS3	MS3	MS3
						MS5	MS5	MS5
							MS4	MS4
								MS2
Simpson's D from %Amp	**0.683**	**0.900**	**0.972**	**0.991**	**0.996**	**0.998**	**>0.999**	**>0.999**
	MS1	MS1	MS1	MS1	MS1	MS1	MS1	MS1
		MS2	MS2	MS2	MS2	MS2	MS2	MS2
			MS7	MS7	MS7	MS7	MS7	MS7
				MS5	MS5	MS5	MS5	MS5
					MS3	MS3	MS3	MS3
						MS4	MS4	MS4
							MS6	MS6
								MS8

Table S3. Simpson's Index of Diversity (D) for combinations of microsatellite markers in order of decreasing heterozygosity (H_e, upper panel) and decreasing marker robustness (%Amp, lower panel). MS6 had the highest heterozygosity (see **Figure 1**, **Table S1**), MS8 had the second highest heterozygosity, and so on. MS1 had the highest PCR and capillary electrophoresis success rate (**Table S1**), MS7 had the second highest, and so on. Two recommended multilocus genotyping schemes are shaded (grey) that achieved high resolution between isolates in our sample set (Simpson's D≥0.999) with the minimal number of loci.

Isolate	MS1	MS3	MS5	MS6	MS7	MS8	Date
SF45-1	**9**	**9**	**2**	**8**	**11**	**11**	7/05/11*
SF45-2	**9**	**9**	**2**	**8**	**11**	**11**	7/05/11*
SF46-1	**9**	**9**	**12**	**12**	**13**	**9**	04/01/10
SF46-2	**9**	**9**	**12**	**12**	**13**	**9**	04/16/10
SF47-1	**10**	**15**	**10**	**11**	**10**	**6**	04/20/09
SF47-2	**10**	**8**	**10**	**11**	**10**	**6**	05/13/09
SF48-1	**5**	**9**	**3**	**9**	**10**	**0†**	08/06/07
SF48-2	**5**	**9**	**3**	**9**	**9**	**0†**	10/10/07
SF49-1	**9**	**16**	**11**	**14**	**10**	**12**	10/25/05
SF49-2	**9**	**0†**	**11**	**12**	**10**	**0†**	03/20/06

Table S4. Paired sample genotypes. Integers represent the absolute number of repeating di- or tri-nucleotide microsatellite units. 1 sample pair (SF45-#) was collected on the same date from different lung lobes of the same patient, while 4 sample pairs were collected from the same patient on different dates. *Samples collected from the same patient on the same date were from different lung lobes (RUL and LLL, respectively). †A value of zero indicates missing data.

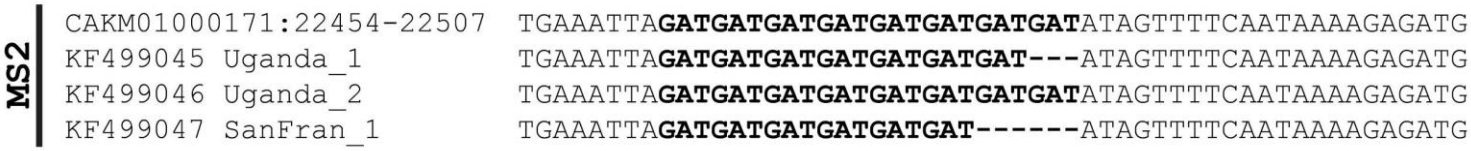

Figure S1. Partial sequence alignments of a microsatellite in a predicted coding region. One microsatellite (MS2) falls within the predicted coding region of the RING-finger membrane protein. Three isolates from Uganda and San Francisco were Sanger sequenced and compared to the reference contig sequence to confirm that the locus is polymorphic in its repeat number.