

Supplemental Data

**Detecting Heteroplasmy from High-Throughput Sequencing  
of Complete Human Mitochondrial DNA Genomes**

Mingkun Li, Anna Schoenberg, Michael Schaefer, Roland Schroeder, Ivane Nasidze, and Mark Stoneking

**Table S1. Primers Used in the SnaPshot™ Validation**

Position	Primers		
	Forward	Reverse	Extension
203	GTA TGC ACG CGA TAG CAT TG	CTG TGT GGA AAG TGG CTG TG	AGG CGA ACA TAC TTA CTA AA
8551	CAC CTA CCT CCC TCA CCA AA	GGG ATC AAT AGA GGG GGA AA	ACG AAA ATC TGT TCG CTT CA
8743	TAT TGA TCC CCA CCT CCA AA	TCC GAG GAG GTT AGT TGT GG	GAA CCT GAT CTC TTA TAC TA
11692	CCA TCT GCC TAC GAC AAA CA	GCG TTC GTA TGA GTT TGC	ACC CCC TGA AGC TTC ACC GG
11809	CAC GGG CTT ACA TCC TCA TT	GGG GGT AAG GCG AGG TTA G	CTC TCT CAA GGA CTT CAA AC
13359	ACA TCT GTA CCC ACG CCT TC	GCT AAT GCT AGG CTG CCA AT	TTC AAA GCC ATA CTA TTT AT
14561	CAA CCA TCA TTC CCC CTA AA	TCC GTG CGA GAA TAA TGA TG	TCA GAA TAA TAA CAC ACC CG
14770	CCC CAC AAA CCC CAT TAC TA	ATC ATG CGG AGA TGT TGG AT	ATC CCA ATA CGC AAA ATT AA
15046	GCC ATG CAC TAC TCA CCA GA	ATG CCG ATG TTT CAG GTT TC	CTC TTC CTA CAC ATC GGG CG



<5%	8	3	0	0	0	3	0	0	0	0
Coverage=72										
>50%	5	1	1	1	1	1	1	1	1	1
25%-50%	2	2	0	0	0	0	0	0	0	0
10%-24%	1	1	0	0	0	0	0	0	0	0
5%-9%	3	1	0	0	0	1	0	0	0	0
<5%	8	3	1	0	0	4	0	0	0	0
Coverage=74										
>50%	5	1	1	1	1	1	1	1	1	1
25%-50%	2	2	0	0	0	0	0	0	0	0
10%-24%	4	3	0	0	0	0	0	0	0	0
5%-9%	3	1	0	0	0	0	0	0	0	0
<5%	11	1	1	0	0	3	0	0	0	0

---

**Table S3. Positions Having Minor Allele Frequency Differences Greater than 0.1 Between Replicates**

Individual	Position	First replicate				Second replicate			
		Major allele		Minor allele		Major allele		Minor allele	
Az4	66	G	70	-	1	G	72	-	24
Az4	309	T	13	C	8	C	68		
Az4	311	C	22	T	5	C	66	T	2
Az4	2457	A	71	-	1	A	76	-	13
Az4	3006	T	25	C	7	T	50		
Az4	3008	C	21	T	10	C	56		
Az4	4715	C	41			C	63	A	7
Az4	11089	A	35			A	42	C	5
Az4	12417	A	68	-	3	A	82	-	26
Az4	16077	A	42			A	62	C	7
Arm3	2458	A	66	-	3	A	74	-	16
Arm3	4774	T	51	A	1	T	42	A	13
Arm3	6820	A	45	C	1	A	50	C	7
Arm3	9314	A	47			A	51	C	7
Arm3	10307	A	36	C	2	A	49	C	9
Arm3	10619	A	60			A	57	C	9
Arm3	10793	A	53	C	1	A	49	C	9
Arm3	12419	A	66	-	11	A	87	-	31
Arm3	15767	A	49	G	1	A	55	C	9
G65	257	A	46	C	8	A	41	C	2
G65	303	C	49	A	1	C	49	-	7
G65	309	C	33	T	8	C	54		
G65	2804	A	62	C	11	A	82		
G65	4773	T	72			T	50	A	6
G65	6691	A	72	-	1	A	72	-	10
G65	10126	A	65			A	56	C	9
G65	10305	A	53	C	4	A	46	C	12
G65	10791	A	72			A	53	C	16
G65	11081	T	56	A	9	T	51		
G65	12417	A	69	-	3	A	83	-	27
G65	14689	A	71			A	60	C	7
G65	15765	A	75			A	49	C	12
Az41	309	C	39	T	9	C	51		
Az41	385	A	60	G	15	A	66	G	6
Az41	4715	C	81	A	2	C	60	A	10
Az41	4773	T	75	A	1	T	43	A	9
Az41	6418	A	83	C	1	A	56	C	8
Az41	7029	A	78			A	55	C	8
Az41	10791	A	81			A	48	C	13
Az41	11081	T	62	A	8	T	50		
Az41	12417	A	79	-	9	A	81	-	21
Az41	14689	A	75			A	57	C	5

**Table S4. Number of Heterozygous Positions in the PhiX174 Sequence Reads, under Different Criteria**

	No quality filter	QS $\geq$ 20	QS $\geq$ 23
Validated by core and ends of the reads <sup>a</sup>	40.3 $\pm$ 5.2 <sup>b</sup>	5.24 $\pm$ 1.8	2.37 $\pm$ 1.15
Validated by ends of the reads	275.4 $\pm$ 38.3	76.7 $\pm$ 8.1	47.9 $\pm$ 11.3
Validated by core of the reads	267.2 $\pm$ 18.0	106.1 $\pm$ 11.3	74.5 $\pm$ 15.8

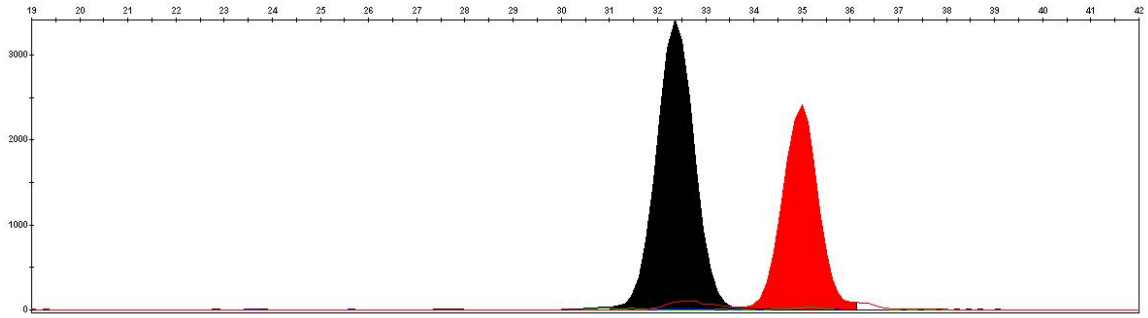
<sup>a</sup> Core of the reads is defined as the content from 6<sup>th</sup> position to 25<sup>th</sup> position of the read and the rest of the reads are defined as the ends of the read.

<sup>b</sup> SD based on 100 resamplings of the data.

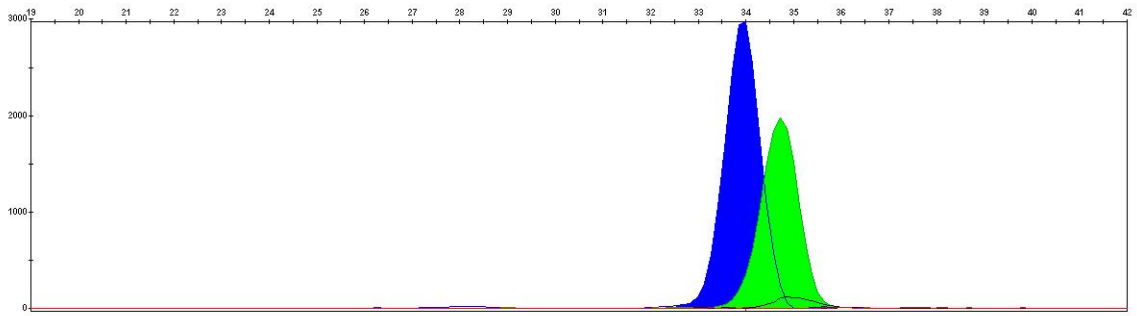
**Table S5. False Positive Error Rate and False Negative Error Rate in Detecting Heteroplasmy by Unique Reads vs. All Reads**

Unique reads				Duplicate reads			
Coverage	False negative rate	False positive rate	Heteroplasmic level	Coverage	False negative rate	False positive rate	Heteroplasmic level
31	0.138	0.060	0.293±0.074	50	0.238	0.067	0.291±0.063
42	0.058	0.011	0.290±0.068	83	0.082	0.048	0.288±0.051
56	0.018	0.0012	0.285±0.059	167	0.0096	0.000801	0.284±0.038
61	0.0076	0.0012	0.281±0.055	250	0.0024	0.0016	0.281±0.031

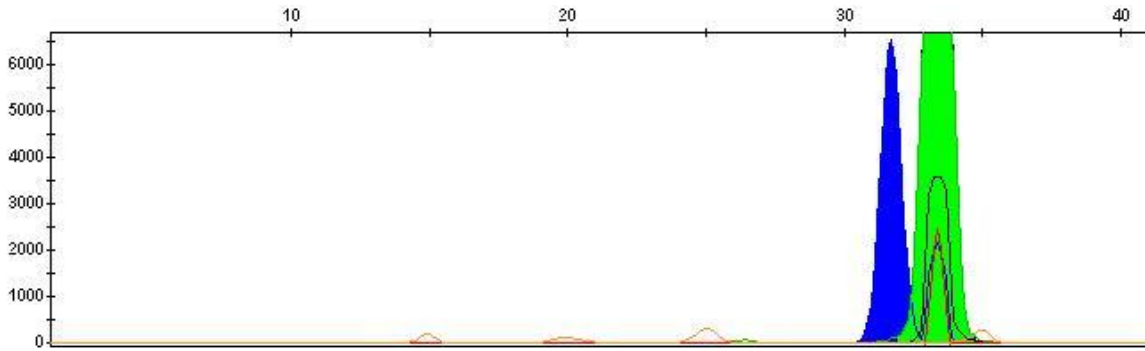
**Figure S1. Electropherograms at the Heteroplasmic Positions by SnapShot™**



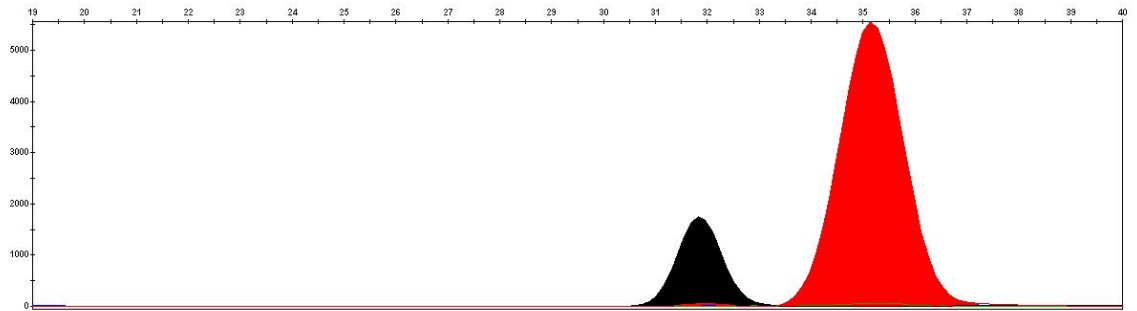
**Position 8551 Ir30**



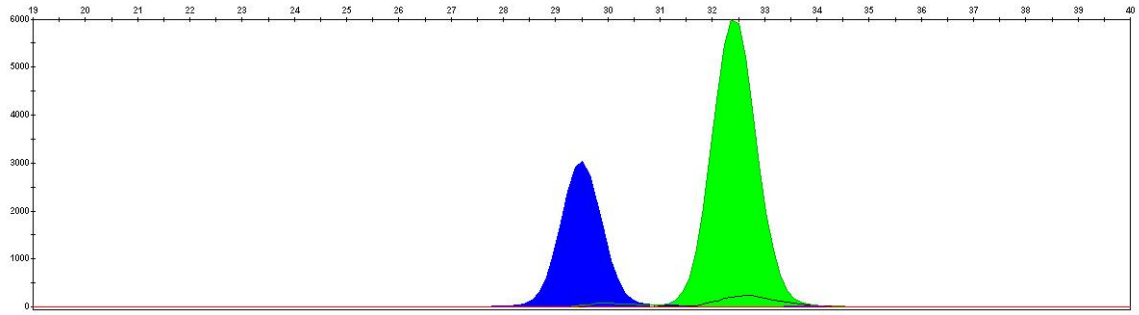
**Position 8743 Ir29**



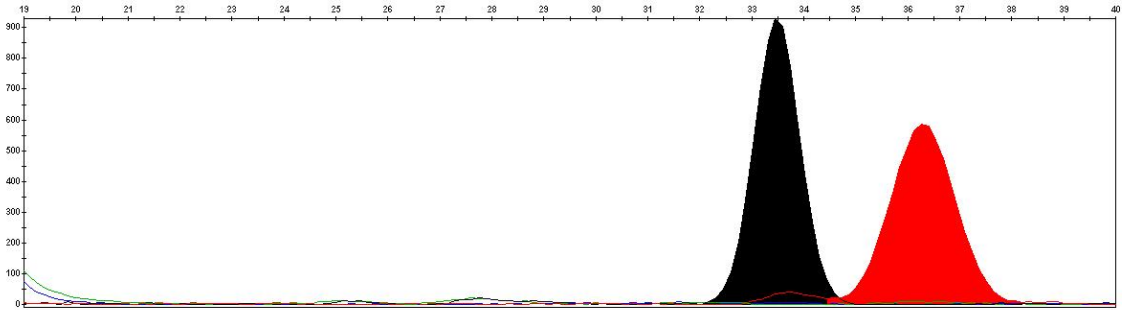
**Position 203 Az5**



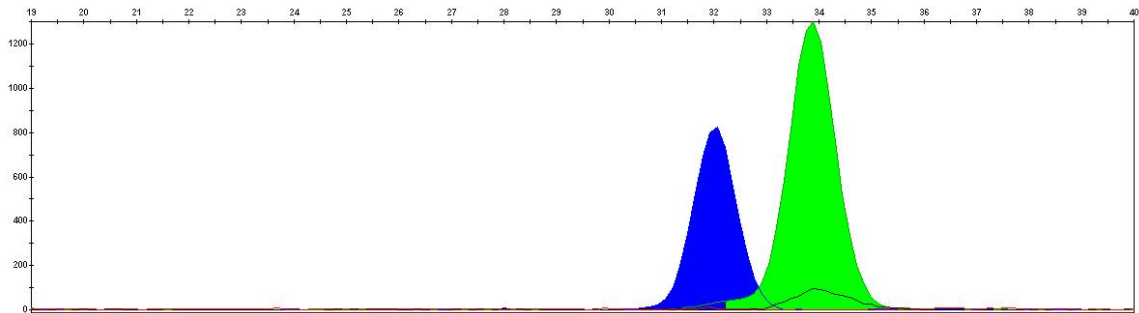
**Position 11809 G82**



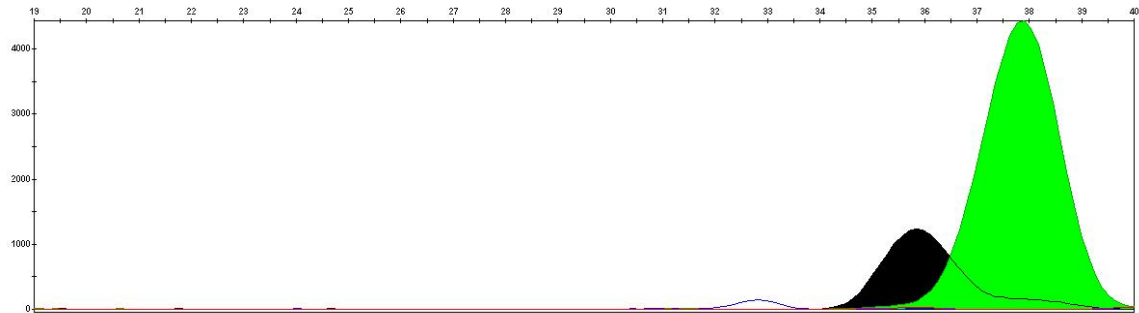
**Position 14561 T9**



**Position 14770 Az14**

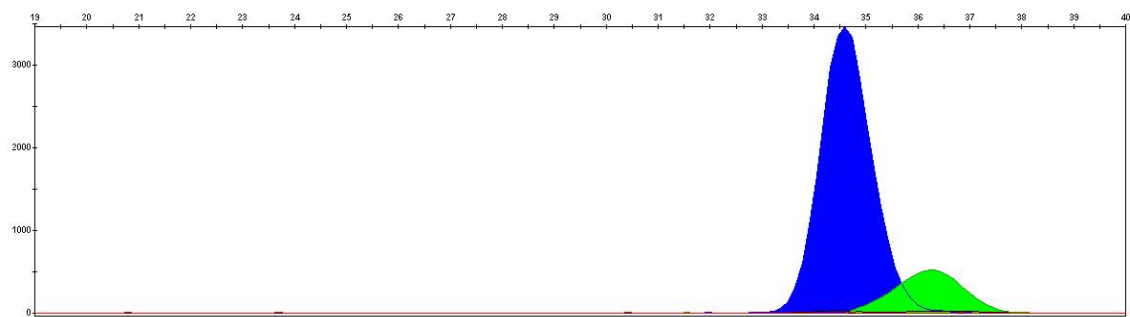


**Position 13359 Az3**



**Position 11692 G38**





**Position 15046 T186**