OMTM, Volume *32*

Supplemental information

mtDNA analysis using Mitopore

Jochen Dobner, Thach Nguyen, Mario Gustavo Pavez-Giani, Lukas Cyganek, Felix Distelmaier, Jean Krutmann, Alessandro Prigione, and Andrea Rossi

Supplemental Information

Supplemental Tables

Table S1: List of primer sequences used in this study.

Target	Forward primer (5'->3')	Reverse primer (5'->3')	Amplicon Size	Task
m14898:151	TAGCCATGCACTACTCACCAGA	GGATGAGGCAGGAATCAAAGAC	1823 bp	Mitopore PCR #1
m16488:1677	CTGTATCCGACATCTGGTTCCT	GTTTAGCTCAGAGCGGTCAAGT	1759 bp	Mitopore PCR #2
m1404:3947	ACTTAAGGGTCGAAGGTGGATT	TCGATGTTGAAGCCTGAGACTA	2544 bp	Mitopore PCR #3
m3734:6739	AAGTCACCCTAGCCATCATTCTA	GATATCATAGCTCAGACCATACC	3006 bp	Mitopore PCR #4
m6511:9220	CTGCTGGCATCACTATACTACTA	GATTGGTGGGTCATTATGTGTTG	2710 bp	Mitopore PCR #5
m8910:10648	CTTACCACAAGGCACACCTACA	GGCACAATATTGGCTAAGAGGG	1739 bp	Mitopore PCR #6
m10360:12226	GTCTGGCCTATGAGTGACTACA	CAGTTCTTGTGAGCTTTCTCGG	1867 bp	Mitopore PCR #7
m11977:13830	CTCCCTCTACATATTTACCACAAC	AAGTCCTAGGAAAGTGACAGCGA	1854 bp	Mitopore PCR #8
m13477:15349	GCAGGAATACCTTTCCTCACAG	GTGCAAGAATAGGAGGTGGAGT	1873 bp	Mitopore PCR #9
m13513	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCTGTACCCACGCCTTC	TGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGCTCAGGCGTTTGTGTATGAT	241 bp	Short read targeted sequencing

Table S2: Calculated number of reads needed to correctly represent a heteroplasmy level of interest with differing confidence intervals.

Heteroplasmy level	Reads for 99% CI	Reads for 95 % CI	Reads for 90 % CI
0.05	15,614	9,011	6,386
0.2	3,287	1,897	1,344
0.35	1,526	881	624
0.5	822	474	336
0.65	442	255	181
0.8	205	119	84
0.95	43	25	18

CI: confidence interval

Table S3: Short-read next generation sequencing results of heteroplasmic samples

ID	Sample	Reads	m13513G>A heteroplasmy level	mtDNA cycles	Amplicon cycles	Barcoding cycles
A1	UMGi176-A cl. 1	951 *	3 %	10	15	19
A2	UMGi176-A cl. 5	367 *	13 %	10	15	19
A3	UMGi176-A cl. 2	260 *	28 %	10	15	19
A4	UMGi176-A cl. 1	18 697	0 %	15	15	19
A5	UMGi176-A cl. 5	1 857	18 %	15	15	19
A6	UMGi176-A cl. 2	1 659	35 %	15	15	19
A7	UMGi176-A cl. 1	15 862	0 %	10	20	19
A8	UMGi176-A cl. 5	2 815	19 %	10	20	19
A9	UMGi176-A cl. 2	2 176	36 %	10	20	19
A10	UMGi176-A cl. 1	91 215	0 %	15	20	19
A11	UMGi176-A cl. 5	35 523	18.5 %	15	20	19
A12	UMGi176-A cl. 2	25 209	38 %	15	20	19

B1	UMGi176-A cl. 1	43 514	0 %	10	25	19
B2	UMGi176-A cl. 5	51 891	20 %	10	25	19
B3	UMGi176-A cl. 2	36 216	39 %	10	25	19
B4	UMGi176-A cl. 1	46 351	0 %	15	25	19
B5	UMGi176-A cl. 5	38 001	19 %	15	25	19
B6	UMGi176-A cl. 2	46 434	38 %	15	25	19

* Below threshold (1000 reads).

Table S4: Comparison of widely used Sanger Sequencing and the Mitopore.de workflow for analysis of mitochondrial DNA samples.

Method	***Time:*** - Hands-on - *To Results*	***Read length (individual amplicon)***	***Recommended DNA input***	***Price per sample***
Sanger Sequencing	- ≤ 1 h - Next working day (if working day is a weekday)	~ 1 kbp	360 ng per sample per amplicon	5 € per amplicon, 80 € for whole mtDNA
Mitopore.de Workflow	- ≤ 6 h - Next working day	~ 2kbp	62.5 ng per sample for all 9 amplicons	19 – 22 € for whole mtDNA

Supplemental Figures

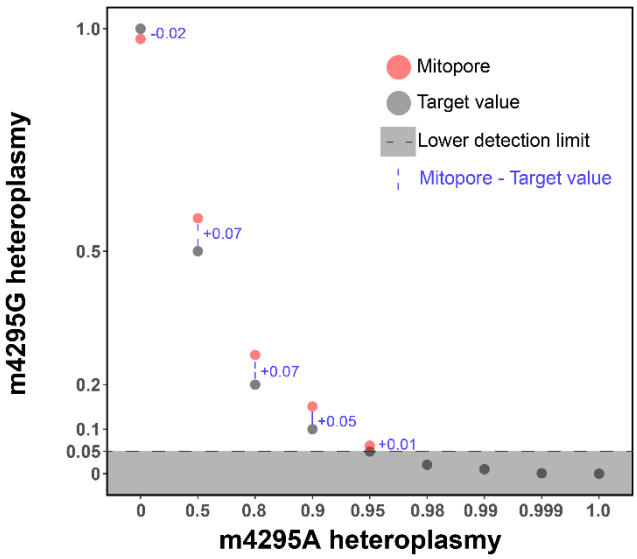

Figure S1: Assessment of the sensitivity of the Mitopore workflow for mitochondrial DNA analysis. DNA from samples with known homoplasmy at position m4295 (G or A) was mixed and the target mitochondrial DNA region amplified and long-read nanopore-sequenced on a Flongle Flow cell. This enables a detection of down to 0.05 heteroplasmy within a range of -0.02 to +0.07 deviation from the real value as determined prior PCR enrichment of the mitochondrial DNA and sequencing.

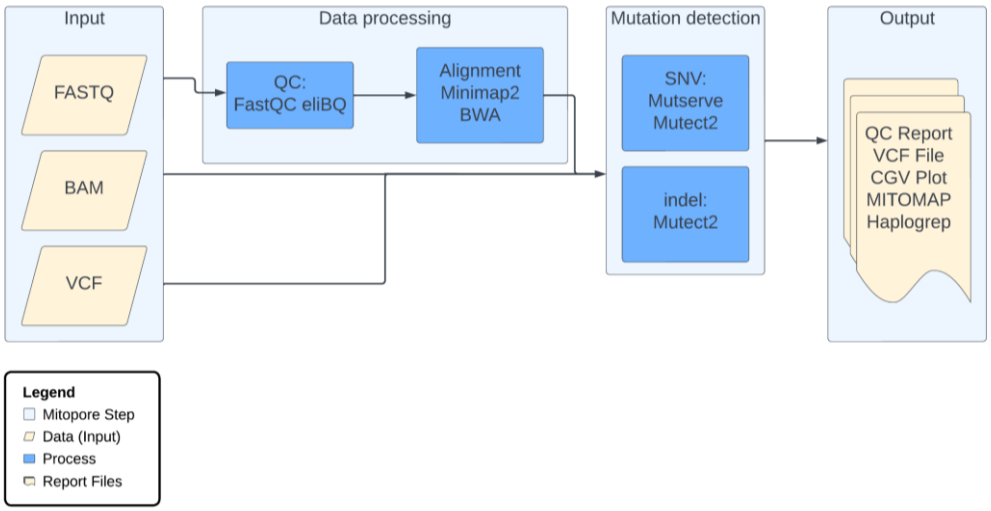

Figure S2: The Mitopore web server workflow. The Mitopore web server workflow consists of three distinct steps after initial data upload by the user: 1.) Data processing including quality control and filtering, and alignment step to the reference genome. 2.) Variant detection including single nucleotide variation/indel calling. 3.) Final report output and export of the results including QC Report, variant calling file (vcf), Circular Genome Visualization (CGV) plot, Mitomap disease annotation and haplotype identification by Haplogrep. All functions are integrated in the Django web server accessible via www.mitopore.de. Flow chart was created in *Lucidchart*, www.*lucidchart*.com.

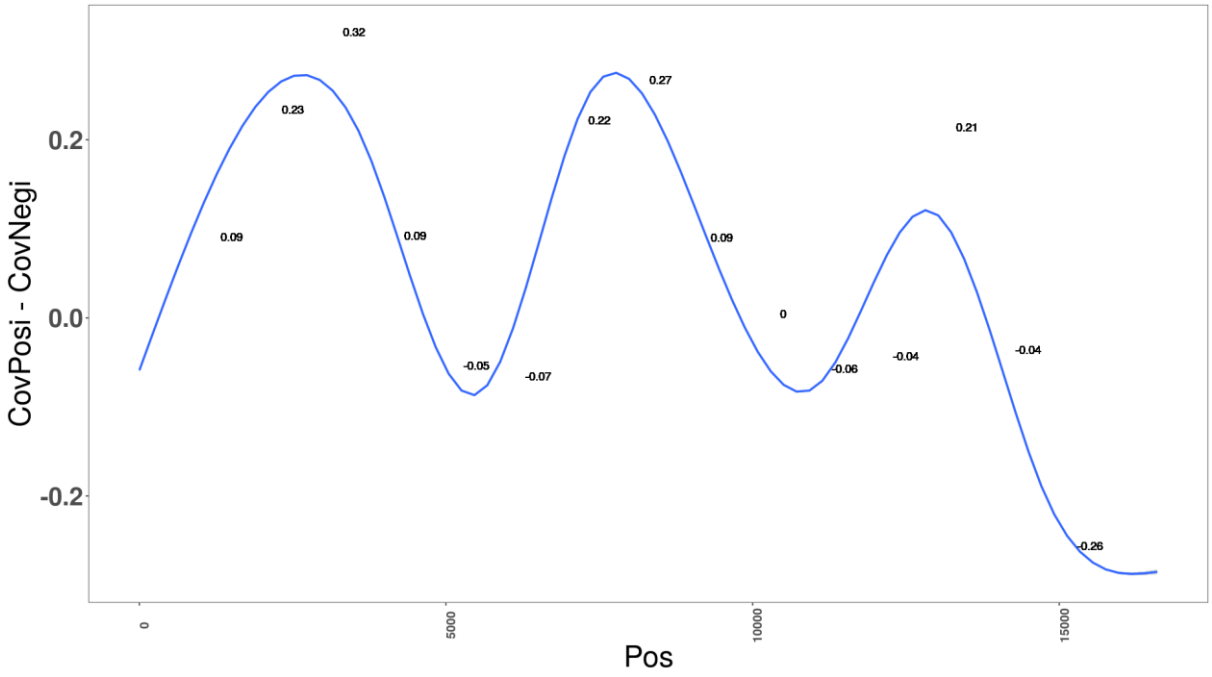

Figure S3: Different coverage between forward and reverse strand. The maximum difference in differential strand coverage of an exemplary sample showing high occurence of variants flagged with STRAND_BIAS is plotted. The maximum difference is 32 % points, indicating uniform strand coverage.

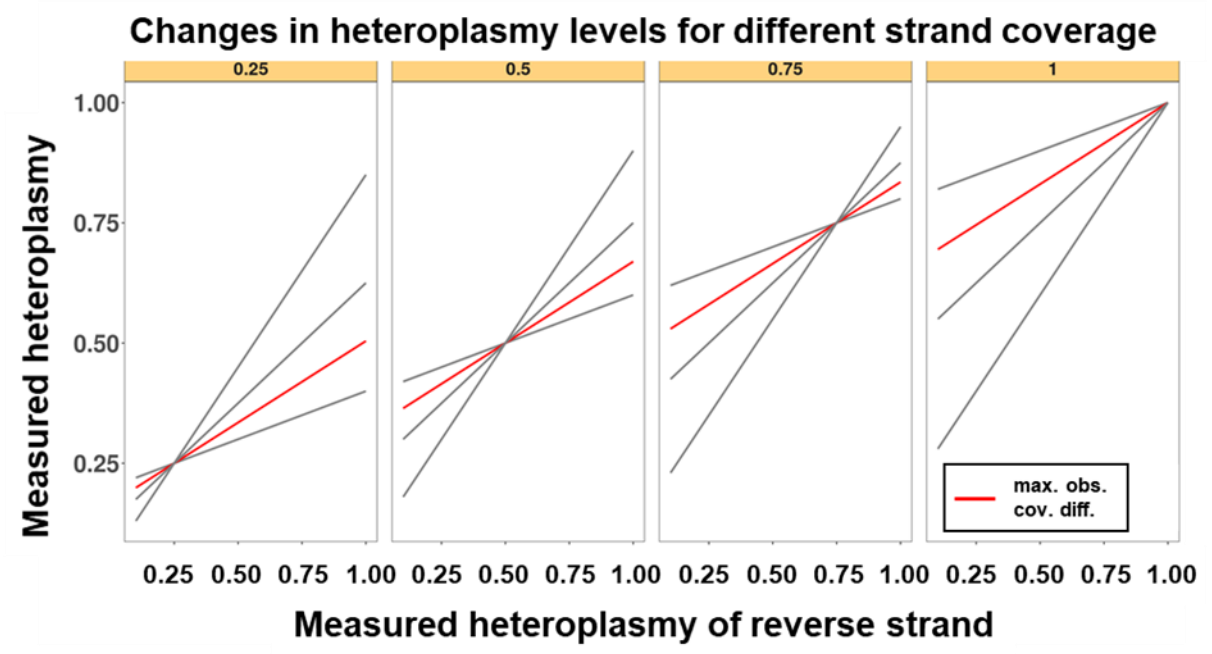

Figure S4: Maximum effects of potentially strand biased coverage on fixed allele frequencies. Individual facets represent (artificial) detected fixed (0.25, 0.5, 0.75, 1)

heteroplasmies on one strand. The resulting heteroplasmy levels are plotted against the detected heteroplasmy levels on the other strand assuming different levels of strand bias. The level representing the maximally observed difference in coverage level (0.339 vs 0.661) is highlighted in red.

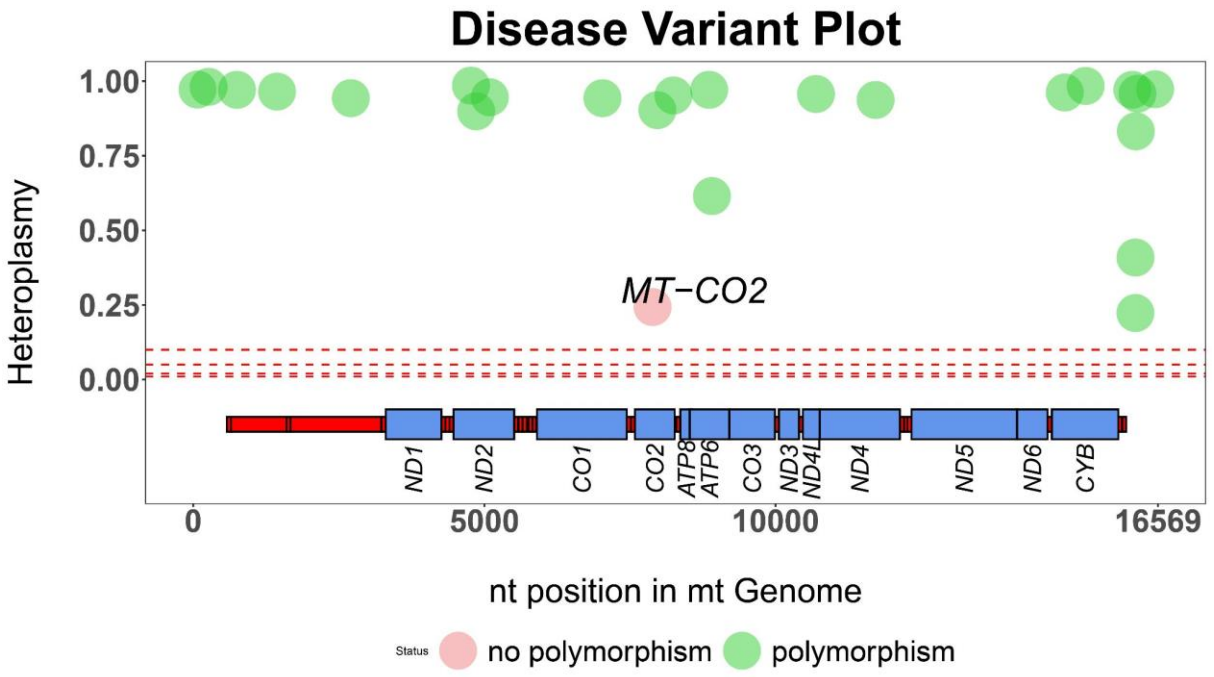

Figure S5: The Mitopore Disease Variant plot enables quick and easy assessment of mitochondrial DNA integrity of analyzed samples. Integrity of mitochondrial DNA (mtDNA) of analyzed samples can be directly evaluated by consulting the Mitopore disease variant plot output. Polymorphic variants (green) are not of concern, while suspicious variants (red) are additionally highlighted by directly indicating the gene of concern.

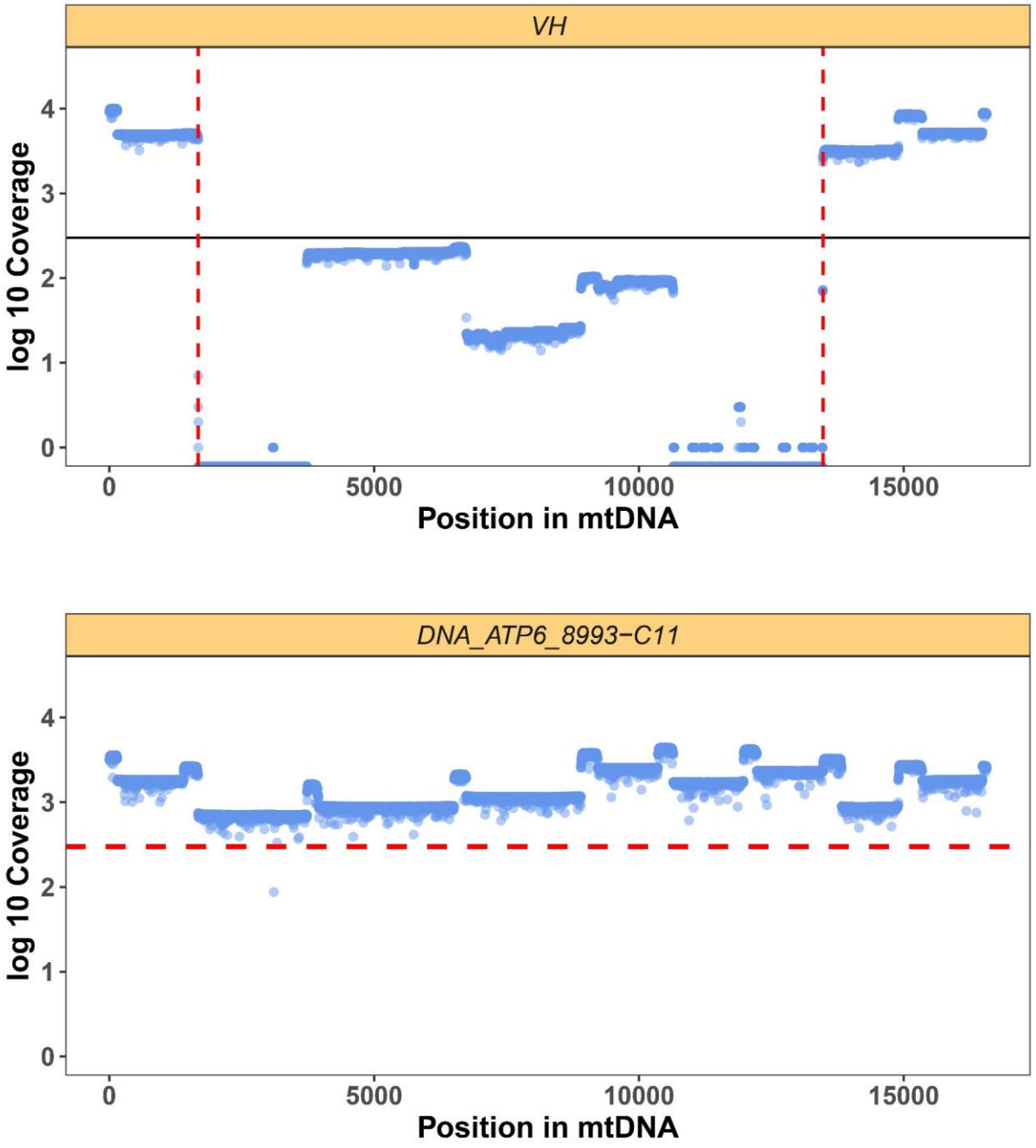

Figure S6: Mitopore enables detection of large deletions. By performing a whole mtDNA PCR enrichment using nine overlapping amplicons, Mitopore enables detection of potential large deletions of several kbp with high accuracy. In the presented example (top, "VH"), initial diagnosis yielded an overestimation of the deletion by roughly 2 kbp. This indicates that the Mitopore workflow and server enable more accurate detection of large deletions associated with mitochondrial diseases. Another example is given for a sample without any indication for a large deletion (bottom, "DNA_ATP6_8993-C11). Each dot represents the number of reads (log10 coverage) at a given position in the mtDNA. The black vertical line indicates the log10

coverage threshold, whereas the red horizontal dashed lines indicate start and end of the identified potential large deletion.

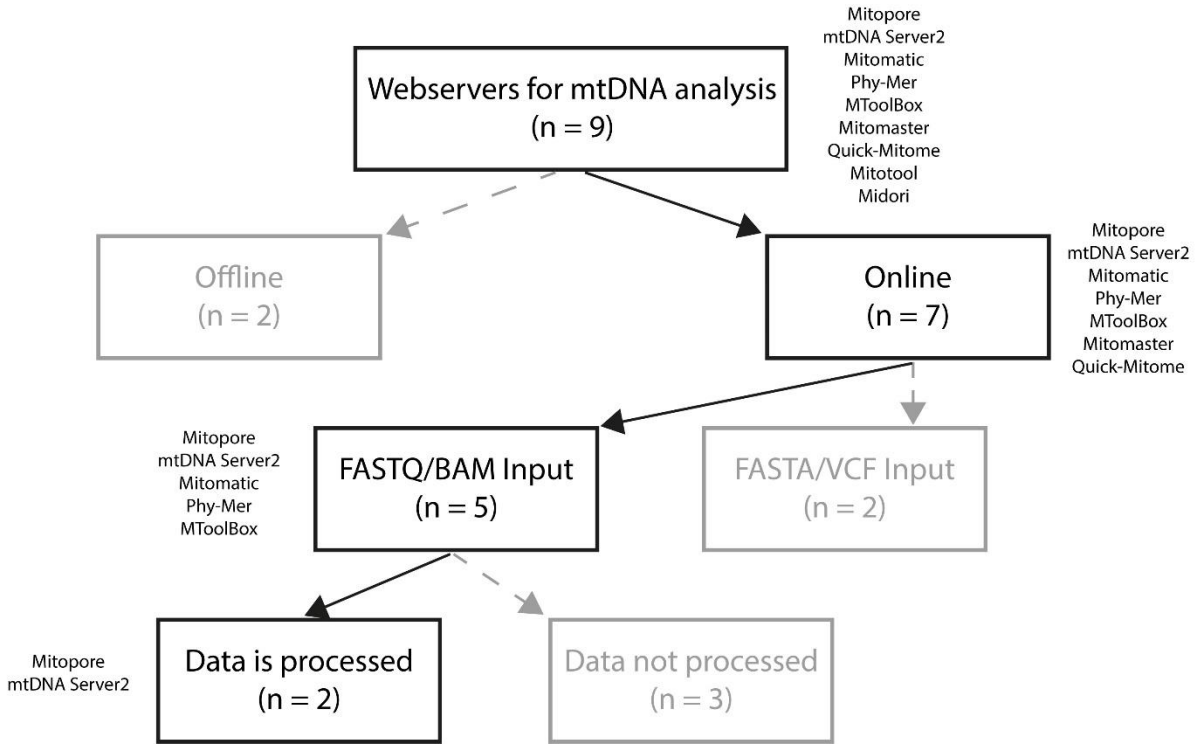

Figure S7: Flowchart showing available webservers for the analysis of mitochondrial DNA. Of the available nine webserver solutions including Mitopore, 29 % were found to be offline. Of the remaining 7, 71 % accept FASTQ or BAM files as input, but 60 % of them were unable to process the data and yield results. A single besides Mitopore one was processed the data, underlining the need for mitochondrial DNA (mtDNA) analysis web server solutions which can handle data from entry level nanopore sequencing devices.

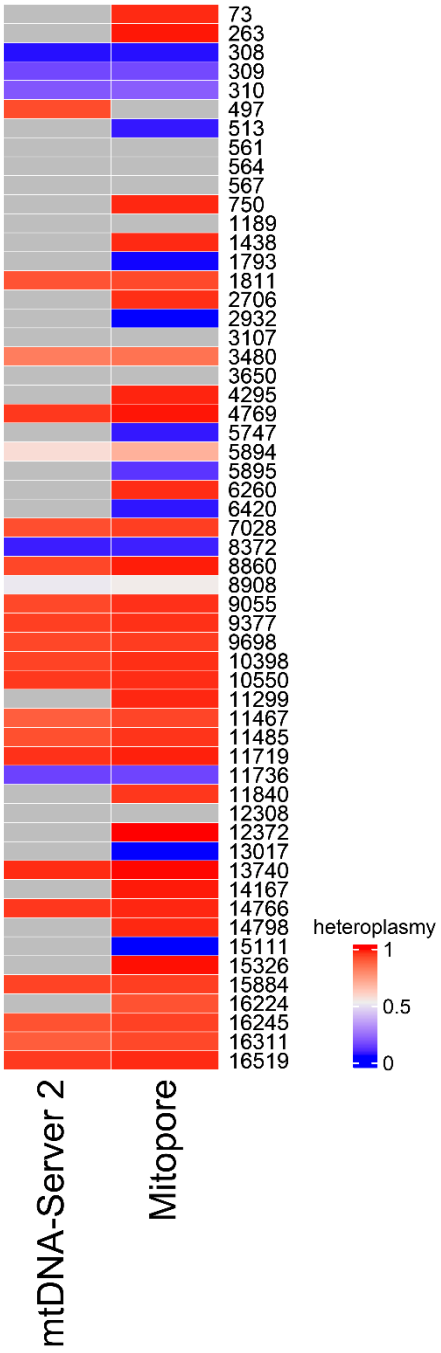

Figure S8: Performance of pipelines for the analysis of mitochondrial DNA derived from low quality nanopore sequencing. PCR-enriched mitochondrial DNA was sequenced on a Flongle Flow cell according to the Mitopore workflow and analyzed using mtDNA-Server 2 or Mitopore. Gray: not detected (cutoff detection threshold: 0.05).