Supplementary Figures and data

Supplementary Figure S1

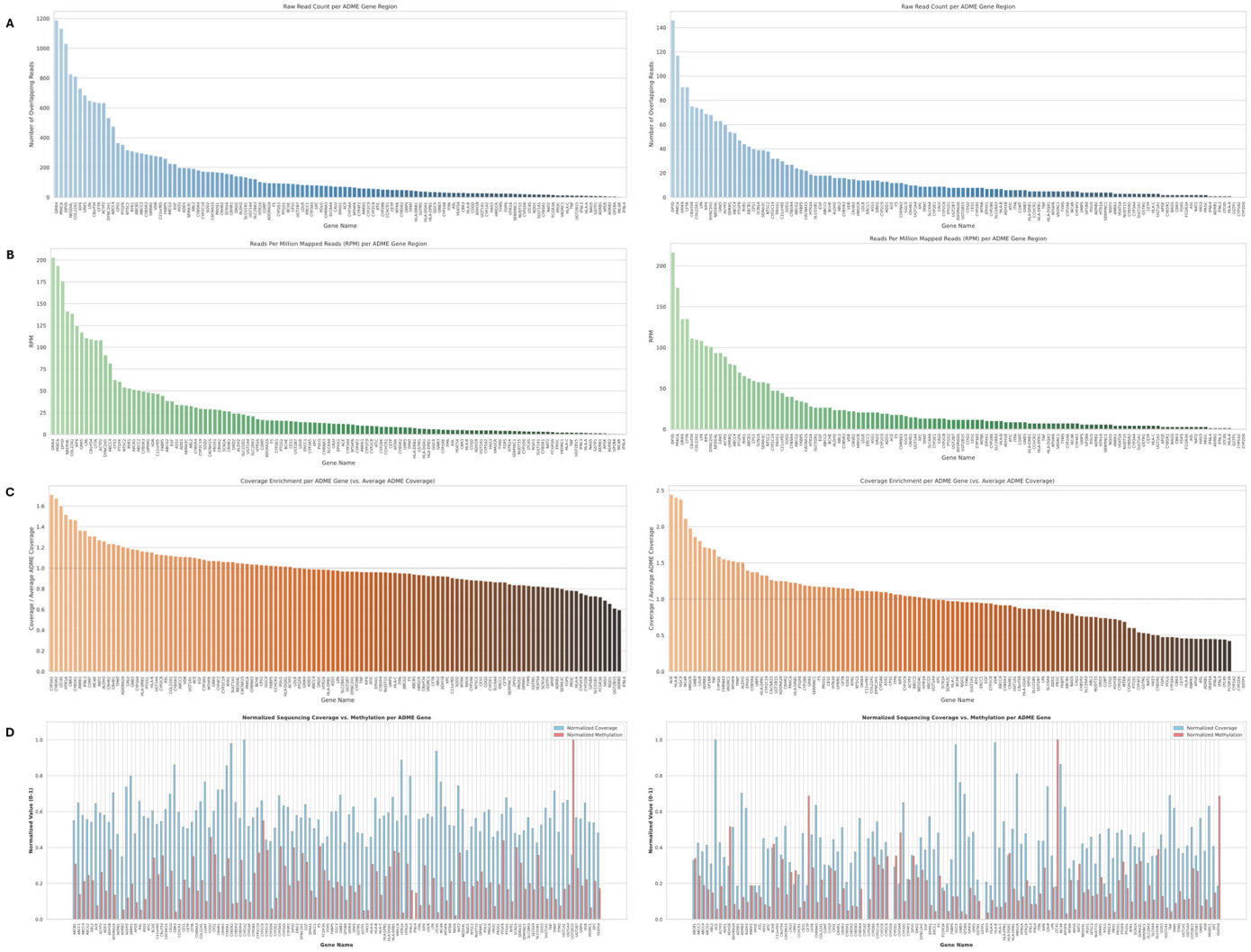


Figure S1. A-C. Comparison of sequencing coverage metrics for selected ADME genes using targeted sequencing (left column) and Whole Genome Sequencing (WGS, right column). Metrics shown include (top row) raw read counts per gene region, (middle row) Reads Per Million mapped reads (RPM) per gene region, and (bottom row) coverage enrichment relative to the average coverage across the analyzed ADME genes within each dataset. D. Normalized sequencing coverage (blue) versus normalized methylation (red) for ADME genes generated by targeted sequencing (left panel) and Whole Genome Sequencing (WGS, right panel).

Supplementary Information: Mitochondrial DNA Sequencing and Haplogroup Analysis

The sequencing data obtained from both adaptive sampling (AS) and whole genome sequencing (WGS) approaches provided detailed information for comprehensive haplotype and haplogroup characterization of the cell lines used in our organ-on-chip models. This characterization offers valuable insights into the genetic background of these cells, which may influence their functional properties and responses in mimicking human tissue environments.

Mitochondrial DNA Variant Analysis

Our analysis revealed distinct mitochondrial DNA profiles for both the JAr and AdMSC cell lines (**Figure S2**). From the mtDNA JAr analysis, we detected 12 variants, 61 heteroplasmies, and 17 inDels. In comparison, the mtDNA AdMSC analysis yielded higher numbers: 34 variants, 30 heteroplasmies, and 42 inDels. This variant profile contributes to the unique haplogroup classification of each cell line, with JAr assigned to haplogroup H7c3 and AdMSC to haplogroup T2b28 (**Figure S3**). These haplogroup assignments were based on specific SNPs and control region mutation motifs identified through our sequencing approaches.

A

Variants and Hetroplasmies

Detected **12 variants**, **61 heteroplasmies** and **17 InDels**.

Variant	Level	Type	Flags
65 (TGG>T)	0.4	InDel	germline;strict_strand;weak_evidence
80 (C>T)	0.07	Heteroplasmy	PASS
263 (A>G)	0.956	Variant	PASS
750 (A>G)	0.974	Variant	PASS
1438 (A>G)	0.957	Variant	PASS
1628 (C>T)	0.041	Heteroplasmy	PASS
1717 (T>C)	0.027	Heteroplasmy	PASS
1748 (G>A)	0.026	Heteroplasmy	PASS
1813 (C>T)	0.034	Heteroplasmy	PASS
1841 (T>C)	0.048	Heteroplasmy	PASS
2061 (C>T)	0.012	Heteroplasmy	PASS
2079 (C>T)	0.046	Heteroplasmy	PASS
2083 (T>A)	0.015	Heteroplasmy	PASS
2135 (A>C)	0.015	Heteroplasmy	PASS
2156 (A>T)	0.018	Heteroplasmy	PASS
2165 (C>T)	0.012	Heteroplasmy	PASS

B

Variants and Hetroplasmies

Detected **34 variants**, **30 heteroplasmies** and **42 InDels**.

Variant	Level	Type	Flags
65 (TG>T)	0.429	InDel	multiallelic
73 (A>G)	0.988	Variant	PASS
263 (A>G)	1	Variant	PASS
308 (C>T)	0.083	Heteroplasmy	PASS
309 (C>T)	0.196	Heteroplasmy	PASS
310 (T>C)	0.241	Heteroplasmy	PASS
709 (G>A)	0.983	Variant	PASS
750 (A>G)	1	Variant	PASS
930 (G>A)	0.965	Variant	PASS
1226 (C>T)	0.127	Heteroplasmy	PASS
1438 (A>G)	1	Variant	PASS
1628 (C>T)	0.059	Heteroplasmy	PASS
1786 (C>T)	0.048	Heteroplasmy	PASS
1888 (G>A)	0.993	Variant	PASS
1924 (T>C)	0.058	Heteroplasmy	PASS

Figure S2: Variants and heteroplasmies summary from Jar **(A)** and from AdMSC **(B)**

A

Summary

Quality Control: PASSED

Haplogroup: H7c3 (0.9683)

Contaminated: NO

Statistics

Parameter	Value
Contig	MT
NumberofReads	3106
CoveredBases	16569
CoveragePercentage	100
MeanDepth	129.611
MeanBaseQuality	18
MeanMapQuality	57.4
RG	@RG,ID:mtDNA_JAR_2,SM:mtDNA_JAR_2

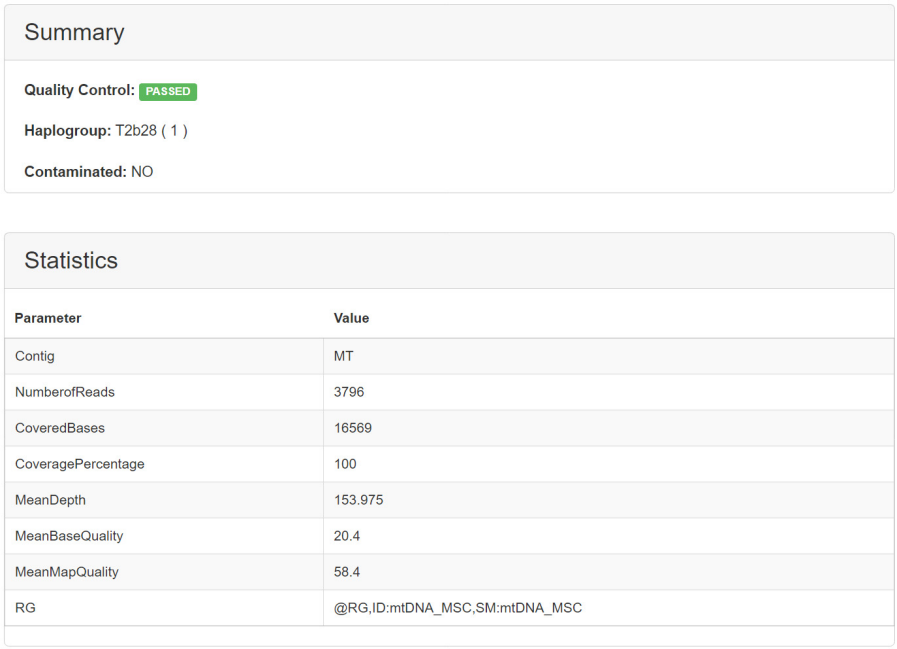


Figure S3: Summary and statistics from the experiments **(A)** for mDNA from Jar **(B)** for mtDNA from AdMSCs

Phylogenetic Analysis

Based on the variant data, we constructed phylogenetic trees for both cell lines (**Figure S4**), revealing their evolutionary relationships and mitochondrial lineages. The phylogenetic analysis demonstrates that both cell lines represent distinct maternal lineages, with characteristic mutation patterns that define their respective haplogroups.

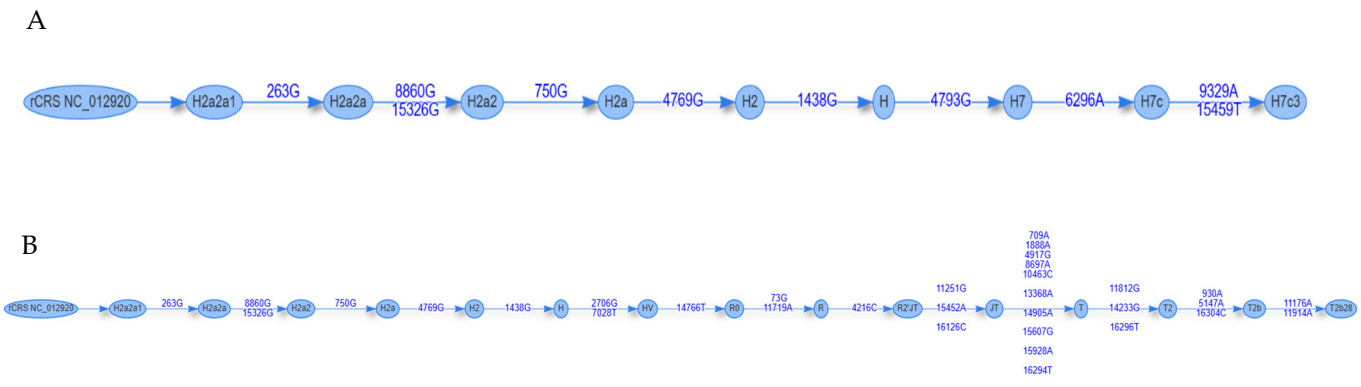


Figure S4: Phylogenetic trees of Jar **(A)** and AdMSC **(B)**

Polymorphism Distribution and Functional Implications

Figure S5 provides a summary of the polymorphisms detected across both cell lines. The dataset includes numerous variants spanning the entire mitochondrial genome, including coding regions, tRNA genes, rRNA genes, and the control region (D-loop). The majority are transitions (purine-to-purine or pyrimidine-to-pyrimidine changes), which are common in mtDNA due to its high mutation rate, with fewer transversions and indels present.

Key variants in the JAr cell line include high-frequency homoplasmic mutations in the control region (D-Loop), such as position 263 A>G (variant level: 0.956) and position 16519 T>C (variant level: 0.985). Additionally, we identified variants in the MT-RNR1 gene (12S rRNA) at position 750 A>G (variant level: 0.974), which can be associated with aminoglycoside-induced deafness in certain contexts, and position 1438 A>G (variant level: 0.957), considered benign. Of particular functional significance is the amino acid change L111P in the MT-ND1 gene (position 3638 T>C), which may disrupt the alpha-helical structure of the NAD-dehydrogenase subunit 1 protein.

The AdMSC cell line exhibited homoplasmic mutations in the control region at positions 263 A>G (variant level: 1.00) and 703 A>G (variant level: 0.988). Significant amino acid changes were observed in several coding regions, including position 3572 T>C and 4097 T>C in MT-ND1, position 13030 T>C (W232R) in MT-ND5, and position 8532 C>T (T56M) in MT-ATP8. These substitutions may affect the function of critical mitochondrial proteins involved in oxidative phosphorylation.

SampleID	Haplogroup	Rank	Quality	Range	Found_Polys	Remaining Polys	Input_Sample
mtDNA_MSC.bam	T2b28	1	1.0000	1-16569	73G 263G 709A 750G 930A 1438G 1888A 2706G 4216C 4769G 4917G 5147A 7028T 8697A 8860G 10463C 11176A 11251G 11719A 11812G 11914A 13368A 14233G 14766T 14905A 15326G 15452A 15607G 15928A 16126C 16294T 16296T 16304C	16519C (hotspot)	73G 263G 709A 750G 930A 1438G 1888A 2706G 4216C 4769G 4917G 5147A 7028T 8697A 8860G 10463C 11176A 11251G 11719A 11812G 11914A 13368A 14233G 14766T 14905A 15326G 15452A 15607G 15928A 16126C 16294T 16296T 16304C 16519C
mtDNA_JAR_2.bam	H7c3	1	0.9683	1-16569	263G 750G 1438G 4769G 4793G 6296A 8860G 9329A 15326G 15459T	6164T (localPrivateMutation) 16519C (hotspot)	263G 750G 1438G 4769G 4793G 6164T 6296A 8860G 9329A 15326G 15459T 16519C

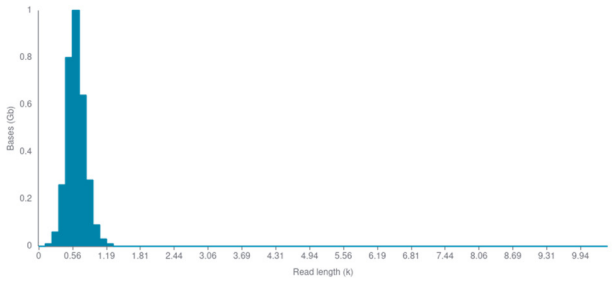
Figure S5: Summary of the polymorphisms. Found_Polys represent the variants that were identified in the sample and matched the expected haplogroup-specific polymorphisms. RemainingPolys represent these that were not found or did not match the expected patterns for the haplogroup, or polymorphisms still to be confirmed.

Connection to Organ-on-Chip Model Characterization

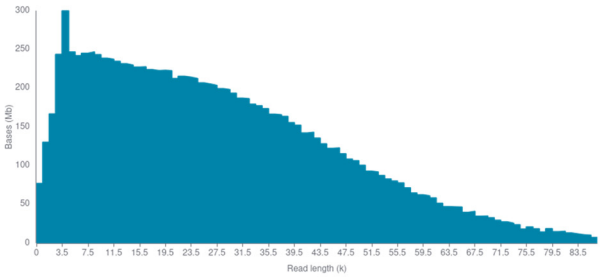
The detailed characterization of mitochondrial haplogroups and variant profiles offers significant benefits for organ-on-chip model development and application. By understanding the mitochondrial genetic background of the cell lines used in these models, researchers can: 1. Account for potential mitochondrial influences on cellular metabolism, energy production, and stress responses. 2. Identify cell lines with specific haplogroups that may better represent certain patient populations or disease conditions. 3. Consider haplogroup-specific drug responses when using these models for pharmacological testing. 4. Establish more precise baseline parameters for normal cellular function in the organ-on-chip systems

Figure S6. Different number of AS and WGS detected reads:

Supplementary: AS + PuvII vs native DNA WGS

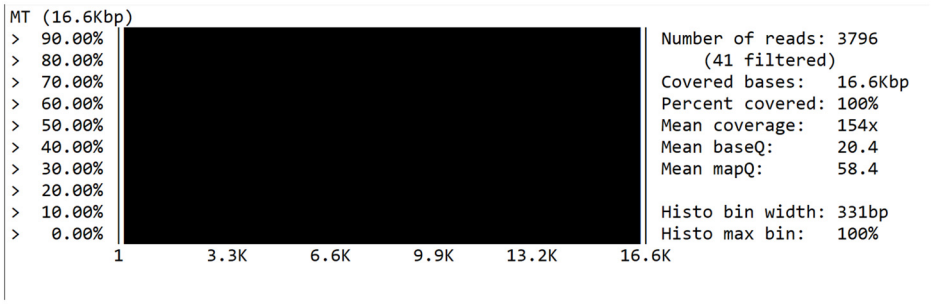


Estimated N50: **701 b**
Estimated bases: **5.69 Gb**
Reads generated: **7.91 M**
Reads called (min Q score: 8): **4.82 M** Pass
Bases called (min Q score: 8): **3.3 Gb** Pass

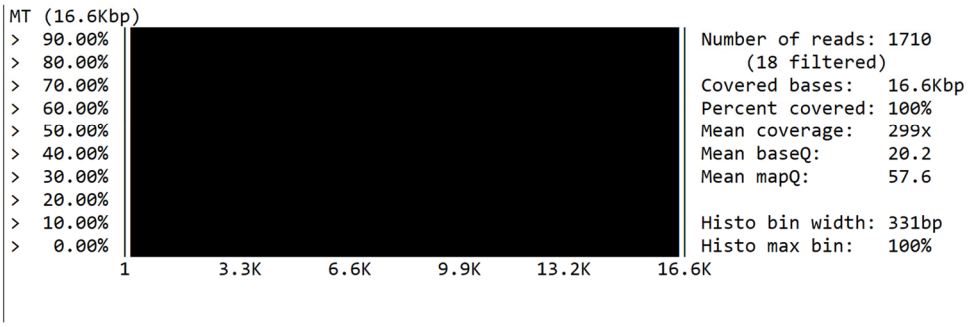


Estimated N50: **25.87 kb**
Estimated bases: **5.69 Gb**
Reads generated:
Reads called (min Q score: 8): **552 036** Pass
Bases called (min Q score: 8): Pass

samtools coverage -r "${MITO_CHROM_NAME}" -m "${INPUT_BAM}"



AS + PuvII mtDNA coverage



Native DNA mtDNA coverage