Single Cell Sequencing of CHO mitochondrial DNA reveals extensive heteroplasmy

Foley. A[12], Clarke. C[12] , Lao. N[1] and Barron. N[12]
[1]National Institute for Bioprocessing Research and Training (NIBRT), Dublin, Ireland
[2]University College Dublin (UCD), Dublin, Ireland
Funded by the STACCATO project as part of the Marie Sklodowska-Curie Action of Horizon 2020 (Grant No. 813453)

Contact:
alan_foley@live.co.uk
niall.barron@nibrt.ie

Introduction

Individual cells within a population can contain wild type and mutant mitochondrial DNA (mtDNA) - a phenomenon known as heteroplasmy. When the proportion of mutant DNA reaches a threshold - phenotypic impacts may be observed.

By analysing each cell individually (single cell), we can view heteroplasmy at a greater resolution than analysing many cells at the same time (bulk).

For example, if bulk analysis identified a variant allele at a frequency of 10% in 10 cells, there are 3 very different possibilities (*Fig. 1*):

(1) All cells contain the allele at a 10% frequency, and are therefore all affected phenotypically
(2) 10% of cells contain the allele at a 100% rate, and therefore only 1 cell is affected phenotypically or
(3) All cells contain the allele at a variable rate (0-100%) and therefore the population will be affected at a variable rate.

Only single cell analysis can differentiate these 3 scenarios

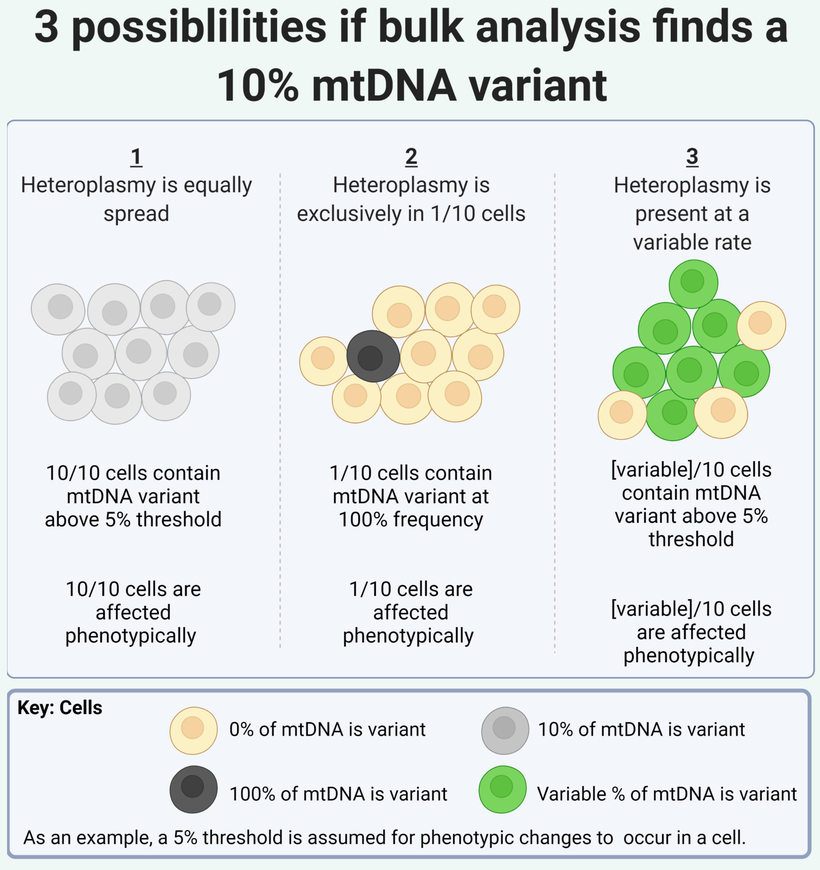

Fig 1: If bulk analysis finds a 10% variant, it can not differentiate between these 3 scenarios; only single cell analysis can

Methods

Four single Chinese Hamster Ovary (CHO) cells, and one bulk sample (4000 cells) were isolated by FACS into lysis buffer.

Lysed contents were purified by AMPure beads. A high fidelity PCR kit amplified the mtDNA by using mtDNA specific preimers. A low PCR cycle number allowed confidence in low frequency heteroplasmies. Samples were run on an agarose gel and mtDNA specific bands gel purified.

Illumina DNA libraries were generated and iSeq100-derived sequencing output was processed and anaylsed using a bespoke bioinformatics pipeline [3].

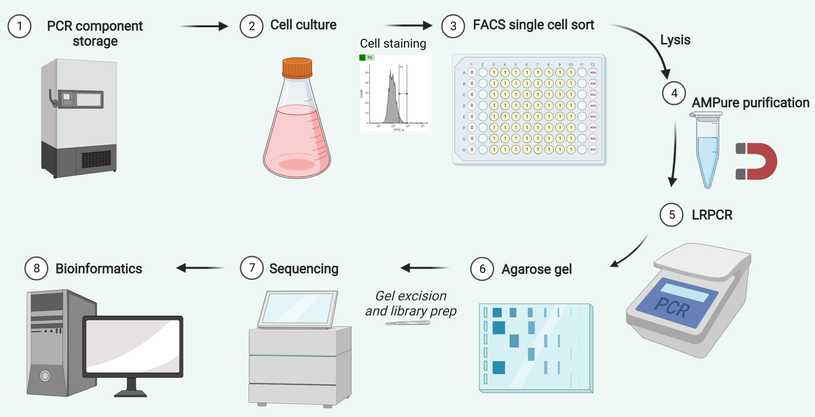

Fig 2: Graphical overview of novel single cell mtDNA platform.
Most steps are designed to cater to the extremely small starting DNA masses

Amplifying mtDNA from a single cell

To test the limits of our PCR protocol, single cell samples were diluted 1:10, 1:100, 1:1000 post-lysis, AMPure purified, and then amplified by a high fidelity Long Range PCR kit (*Fig 3*). Samples were split into two separate reactions to amplify the 16.5kb mtDNA genome in 2 halves (8.5kb). This allowed sequencing coverage of overlapping regions.

An 8.5kb electrophoresis band, indicative of mtDNA, was observed in the positive control of 10 cells and no band in the negative control.

We here present the first evidence of single cell mtDNA amplification in CHO cells with a mtDNA band in the single cell lane.

Furthermore an 8.5kb band was observed even with a 1/1000 dilution. Since there are an estimated 1,000-10,000 mitochondria per cell [2], this method may even be viable for single mitochondrial sequencing.

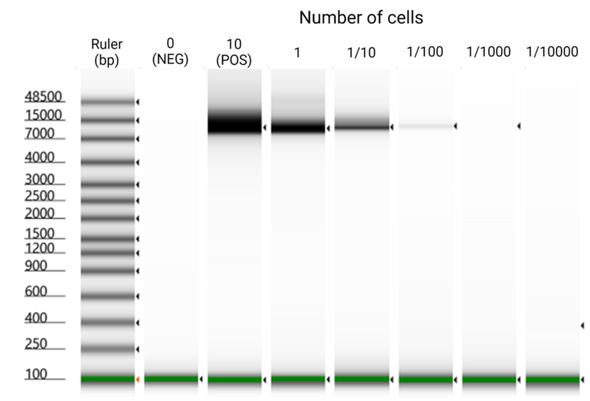

Fig 3: TapeStation automated electrophoresis of PCR amplified mtDNA

Identification of mtDNA variants in individual CHO cells

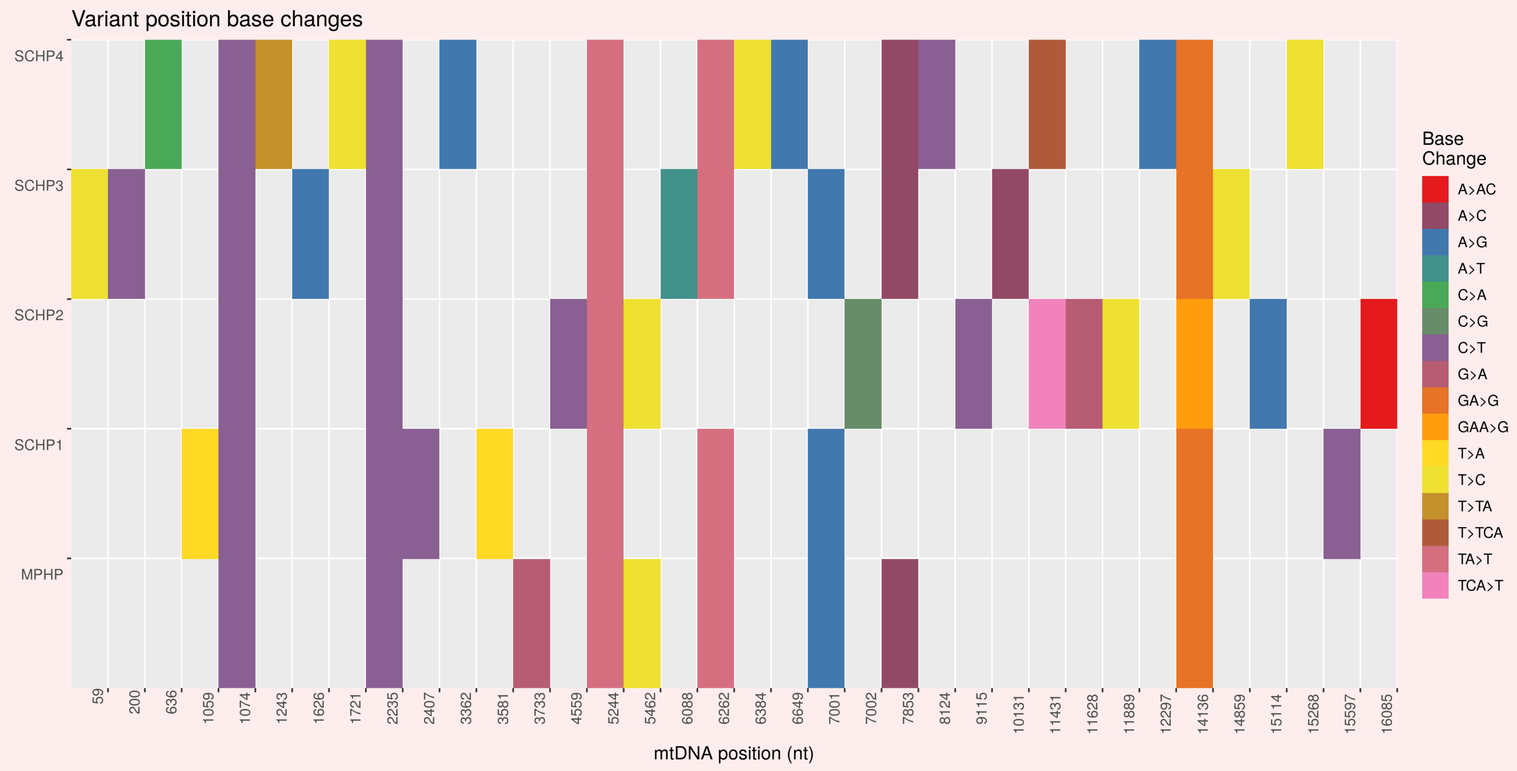

Fig 4: Variant locations in mtDNA. SCHP1 = Single Cell 1, SCHP2 = Single Cell 2, SCHP3 = Single Cell 3, SCHP4 = Single Cell 4, MPHP = 4000 cells
x-axis = position along mtDNA genome, y-axis = sample

After gel purification, sequencing and pre-processing, mtDNA variants were called by LoFreq and VarScan from 4 single cells and 1 mixed population. If both tools called a variant above 2%, it was brought forward for analysis.

The x-axis presents position along the 16.5kb CHO mtDNA genome. The y-axis shows 4 single cells and 1 mixed population of 4000 cells; allowing analysis of base change (Fig. 4) and variant impact (Fig. 5).

75% (27/36) of variant locations were exclusively found among the single cells but not in the mixed population.

Frameshift variants, which are likely to be highly impactful, were observed in protein-coding genes COX1, CYTB and ND4 at positions 6262, 11431 and 14136 (*Fig. 5*).

Conclusions

Fig. 5: Effect of variants in mtDNA as predicted by SNPeff
x-axis = position along mtDNA genome, y-axis = sample

If heteroplasmy is evenly spread among all cells, there should be an equal number of variants called in the bulk sample and the single cells; but this is not observed (*Fig 4+5*).

This supports an uneven spread of heteroplasmy as in scenario 3 (*Fig 1*).

We present the first evidence of cell-to-cell variation in CHO cell mtDNA, indicating a possible contribution to heterogenity in cells; and also revealing possible engineering targets.

References:
[1] Kelly. P et al. (2017). Ultra-deep next generation mitochondrial genome sequencing reveals widespread heteroplasmy in Chinese hamster ovary cells. Metabolic Engineering 41, 11–22. doi:10.1016/j.ymben.2017.02.001.
[2] Dhiman, H et al. (2019). Genetic and Epigenetic Variation across Genes Involved in Energy Metabolism and Mitochondria of Chinese Hamster Ovary Cell Lines. Biotechnology Journal 14, 1800681. doi:10.1002/biot.201800681.
[3] github.com/alanfoleynibrt/SingleCellmtDNA)