Single Cell Whole Genome Amplification

Alison Wakefield[1], Paresh Patel[1], Ryan C. Heller[2], Andrew Gane[1], Stephen Capper[1] & John R. Nelson[2]

[1]GE Healthcare, Cardiff, Wales, UK

[2]GE Global Research, Niskayuna, New York, US

1. Introduction

- The analysis of genomes at the single cell level offers unprecedented biological insights in diverse fields such as cancer research, immunology & microbiology. To enable single cell genomics, a technology for amplification of genomic DNA is required that provides utmost sensitivity, accuracy & robustness.
- We have previously developed a method of multiple strand displacement amplification (MDA) by Phi29 DNA polymerase.

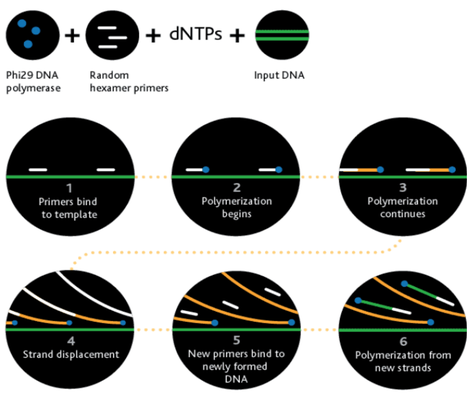


Figure 1: Overview of Multiple Displacement Amplification by Phi29 DNA Polymerase.

- This single subunit, proofreading DNA polymerase has excellent processivity and possesses strand displacement properties that enable the high-fidelity amplification of input DNA using random hexamer primers.
- Phosphorothioate modification of the primers prevents degradation by the DNA polymerase and dramatically stimulates reaction kinetics. This system is commercially available as the illustra™ Single Cell GenomiPhi™ DNA Amplification Kit .
- The project presented here builds on the merits of conventional MDA to develop methods to recover and amplify DNA from single cells of bacterial and mammalian origin. Quality of the output DNA in terms of genome coverage, uniformity of amplification, and error rate is critical to obtain useful single cell data in various applications. Microarray analysis and NGS are used to validate our novel single cell MDA protocols and formulations.

2. Methods

- A number of improved GenomiPhi formulations, named GRC1, GRC2 and CD, were developed for single cell WGA. These formulations were then optimized to develop the Single Cell GenomiPhi product.
- New manufacturing processes, including UV & enzymatic reagent clean-up, help to ensure that all kit components are free from any detectable DNA contamination & enable sensitivity of amplification down to 1 fg of gDNA. An optional, proprietary, enzymatic clean-up step ensures that any potential DNA contaminants introduced during setup are removed before each individual reaction.

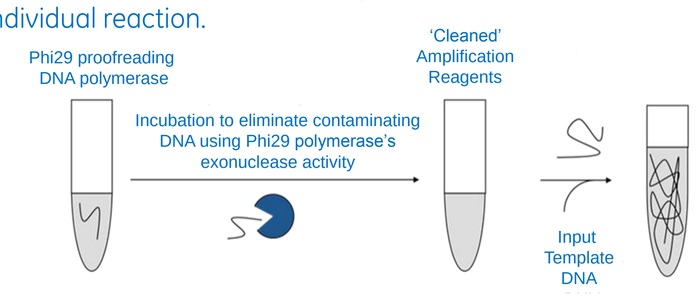


Figure 2: Enzymatic Cleaning Reaction

- New formulations have been tested on both microbial and mammalian cells. Real time WGA kinetics were performed by adding SYBR™ Green I to the amplification mixture and monitoring fluorescence increase over time in a Tecan™ plate reader.
- Amplified DNA quality was assessed by performing downstream sequencing (illumina™ HiSeq™) and SNP analysis (Affymetrix™).

3. Results

A) New Formulation is Sensitive Enough to Amplify from Femtograms of gDNA

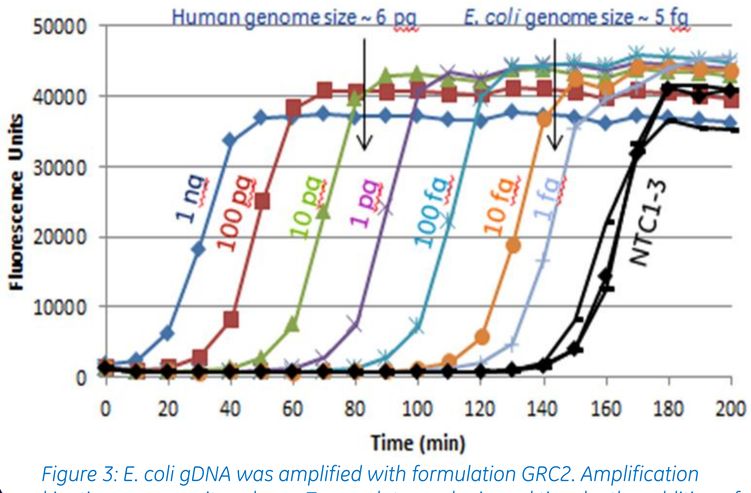
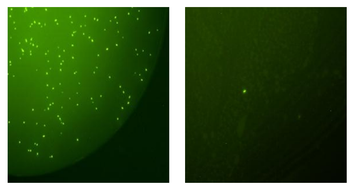


Figure 3: E. coli gDNA was amplified with formulation GRC2. Amplification kinetics were monitored on a Tecan plate reader in real time by the addition of SYBR™ Green I.

B) New Formulations Consistently Amplify from Single Microbes with Supressed Background Amplification

a.



b.

Sample	Conc. (ng/ul)	Yield (ng)
A17	267	6141
B2	267	6141
B5	296	6808
D2	110	2530
D7	22	506
D8	46	1058
J6	158	3634
J8	226	5198
J9	224	5152

c.

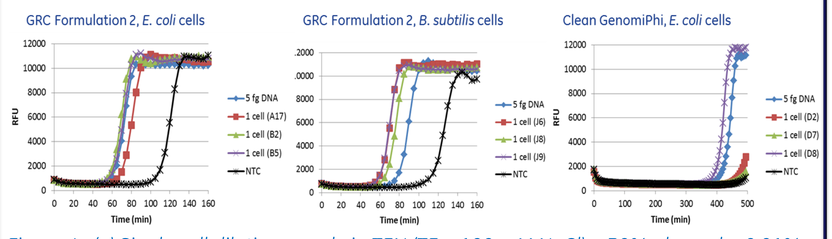


Figure 4: (a) Single cell dilutions made in TEN (TE + 100 mM NaCl) + 30% glycerol + 0.01% Tween-20 with the addition of FM1-43 FX dye for visualization. (b) Single cell post-amplification yields. (c) Real-time amplification kinetics using GRC formulation 2 on E. coli *and* B. subtilis *single cells and clean GenomiPhi (current GenomiPhi V2 formulation + enzymatic cleaning reaction) on* E. coli *single cells.*

C) Modified Cell Lysis Method Improves Genome Coverage of DNA Amplified from Single Microbial Cells

a.

b.

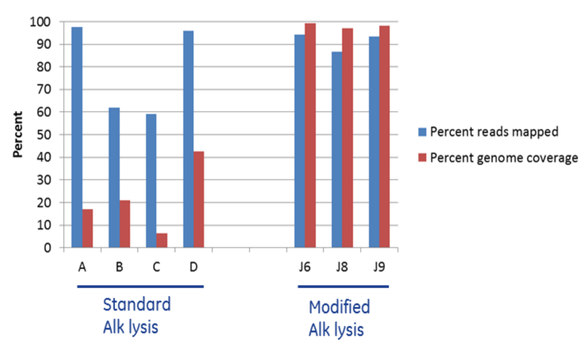


Figure 5: (a) Whole Genome Sequencing (illumina™ HiSeq™) of DNA amplified from single B. subtilis *cells with GRC formulation 2 using two different alkaline lysis methods. (b) New alkaline lysis method increases genome coverage of amplified DNA.*

D) New Formulations Reduce Amplification Bias in DNA Amplified from Single Microbial Cells

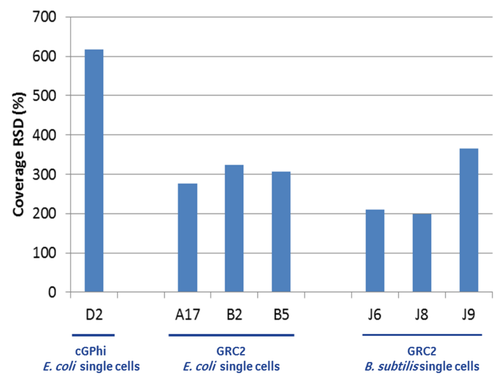


Figure 6: Whole Genome Sequencing (illumina™ HiSeq™) of DNA amplified from single E. coli *and* B. subtilis *cells with GRC formulation 2. RSD = Relative Standard Deviation = Mean Coverage standard deviation/Mean coverage x 100.*

E) Single Cell GenomiPhi Enables Robust Mammalian Cell Lysis and Amplification Kinetics

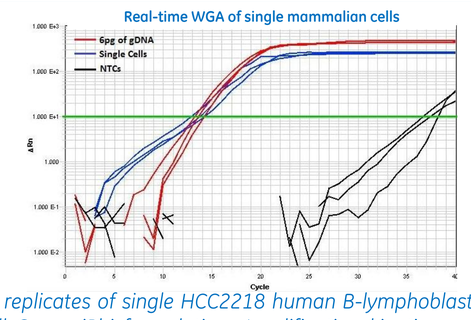
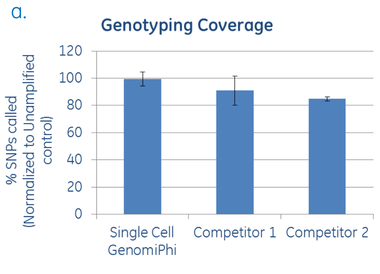


Figure 7: Three replicates of single HCC2218 human B-lymphoblast cells were amplified with Single Cell GenomiPhi formulation. Amplification kinetics were monitored in real time by the addition of SYBR™ Green I (1 cycle = 5 min).

F) Single Cell GenomiPhi Shows Improved Coverage and Accuracy from Mammalian Cells in SNP Analysis

a.



b.

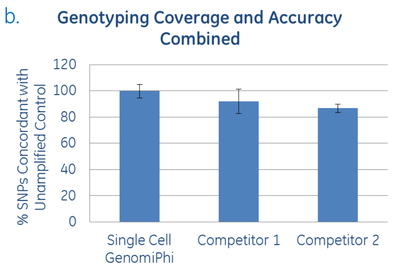
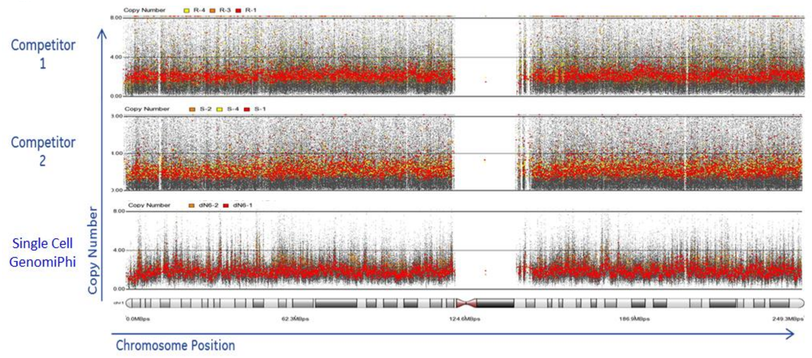


*Figure 8: gDNA amplified from 5 cell Jurkat samples using different WGA technologies was run on an Affymetrix™ Genome Wide Human 6.0 SNP Array and compared to unamplified control gDNA. (a) Genome coverage is represented by % SNPs successfully assigned a genotype. (b) Coverage and accuracy combined is represented by % SNPs with a concordant call to the unamplified Jurkat control (n=2). *See footnote below.*

G) Single Cell GenomiPhi Shows Reduced Amplification Bias from Mammalian Cell Samples Compared to Competitor WGA Techniques

a.



b.

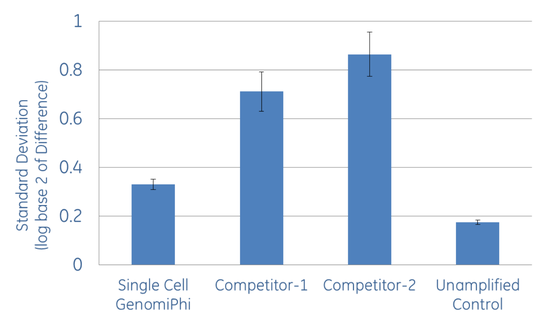


*Figure 9: gDNA amplified from 5 Jurkat cell samples using different WGA technologies was run on an Affymetrix™ Genome Wide Human 6.0 SNP Array and compared to unamplified control gDNA. (a) Copy number graph showing WGA DNA normalised to the unamplified bulk control gDNA samples. Graphs represent overlap of 3 separate samples. Colored dots = mean intensity of 30 probes, Representative graph of chromosome 1 is shown. (b) Graph showing amplification bias of different WGA technologies. Each spot intensity was compared to that same spot on one of the control slides. Standard deviation of each slide compared to that control is shown (n=2). *See footnote below.*

H) Single Cell GenomiPhi Shows Good Coverage from Mammalian Cell Samples in Whole Exome Sequencing

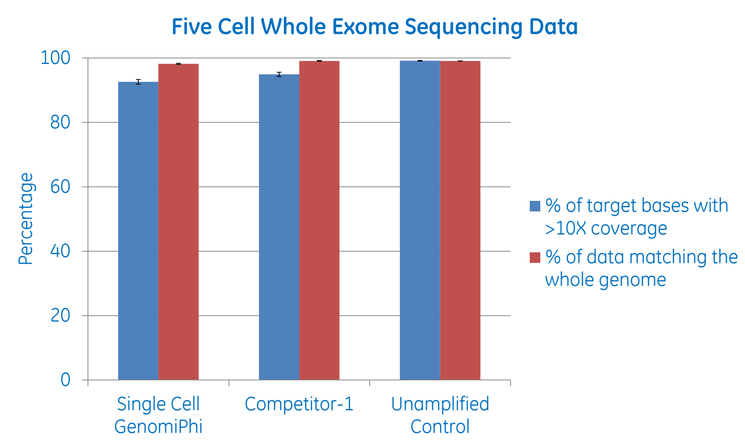


*Figure 10: Five HCC2218 human B-lymphoblast cells were lysed, and the gDNA was amplified using Single Cell GenomiPhi and a competitor WGA kit. SureSelect™ All Exon Target Enrichment System (Human All Exon Kit V5, Agilent™) was used to capture the exomes of amplified test and unamplified control samples. All samples then underwent sequencing on illumina™ HiSeq™ 2500 platform. DNA sequences were aligned to hg19 using BWA-0.6.1. Figure shows the percent of data matching the whole genome and percent exome coverage at 10x read depth (n=2). *See footnote below.*

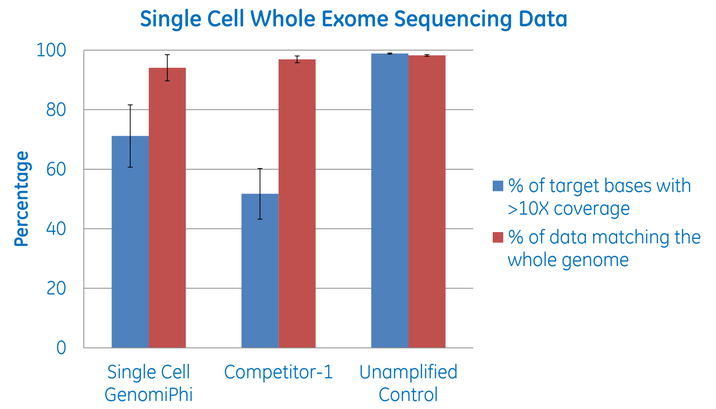


*Figure 11: Single HCC2218 human B-lymphoblast cells were directly dispensed into 96-well qPCR plate containing lysis buffer using Becton Dickinson FACSAria III. gDNA was amplified using Single Cell GenomiPhi and a competitor WGA kit. SureSelect™ All Exon Target Enrichment System (Human All Exon Kit V5, Agilent™) was used to capture the exomes of amplified test and unamplified control samples. All samples then underwent sequencing on illumina™ HiSeq™ 2500 platform. DNA sequences were aligned to hg19 using BWA-0.6.1. Figure shows the percent of data matching the whole genome and percent exome coverage at 10x read depth (n=3). *See footnote below.*

4. Summary

- Single Cell GenomiPhi provides a robust and representative method for whole genome amplification from single cells.
- New optimized formulation increases the amplification sensitivity down to 1fg amount of DNA.
- New manufacturing process has been developed that includes UV decontamination and enzymatic clean-up to eliminate any contaminating DNA from whole genome amplification reagents.
- Single cell whole genome amplification, SNP arrays and the NGS results suggest that the Single Cell GenomiPhi kit provides improved sensitivity and coverage with reduced amplification bias.

Poster presented by Andrew Gane, from GE Healthcare Cardiff on 20-21st August 2014 at the 2nd Annual Single-Cell Sequencing Conference. Some work presented was supported by NIH grant #1R21HG005065-01A1. The opinions, findings, and conclusions or recommendations expressed in this publication/program/exhibition are those of the authors and do not necessarily reflect the views of the National Institutes of Health. All data has been generated and reviewed in compliance with the GE Healthcare quality management system. GE and GE monogram are registered trademarks of General Electric Company. GenomiPhi and illustra are registered trademarks of General Electric Company or one of its subsidiaries. Agilent and SureSelect are registered trademarks of Agilent Technologies. GenomePlex is a registered trademark of Sigma-Aldrich Corp. SYBR is a registered trademark of Life Technologies Corporation. REPLI-g is a registered trademark of Qiagen. Tecan is a registered trademark of Tecan Group Ltd. Affymetrix is a registered trademark of Affymetrix Inc. Illumina and HiSeq are registered trademarks of Illumina, inc. General Electric Company reserves the right, subject to any regulatory approval if required, to make changes in specifications and features shown.

First published Nov. 2013

GE Healthcare UK Limited, Amersham Place, Little Chalfont, Buckinghamshire, HP7 9NA UK

*Preliminary data in figures 8 & 9 are from testing performed between 6th Mar–8th April 13, data in figure 10 is from testing performed between 3rd Oct – 29th Oct 13 and data in figure 11 is from testing performed between 25th Mar-15th July 14 by GE Healthcare. WGA was performed at GE Healthcare, Cardiff and subsequent SNP arrays and NGS were performed at SeqWright, Houston, Texas.
Competitor 1 = Qiagen REPLI-g Single Cell kit, Competitor 2 = Sigma GenomePlex Single Cell Whole Genome Amplification Kit.

To whom all correspondence should be addressed: alison.m.wakefield@ge.com