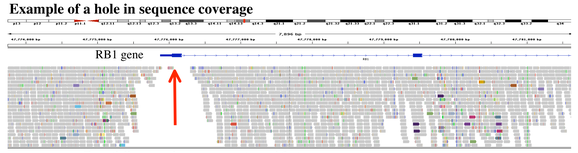
Analysis of Bias during PCR and Whole Genome Amplification

Maxwell, Daniel Aird and Andreas Gnirke
Broad Institute, 320 Charles, Cambridge, MA

Introduction

The GC content of DNA, i.e. the percent guanine and cytosine, is not uniform along the genome. Sections of genomes can be very high, neutral or low in %GC. Amplification technologies, such as Polymerase Chain Reaction (PCR) and Whole Genome Amplification (WGA), are potentially sensitive to base composition. This means that some sections are not copied as well as the more GC neutral sections of the genome. This is a problem when PCR is used to amplify genomic DNA for genomic sequencing. Due to the bias during PCR, specifically in the Illumina library prep protocol, the genome of the organism being sequenced will have holes that are either under represented, incomplete or nonexistent in the sequencing reads.

Example of a hole in sequence coverage



Essentially no reads cover promoter and first exon (76% GC) of this tumor suppressor.

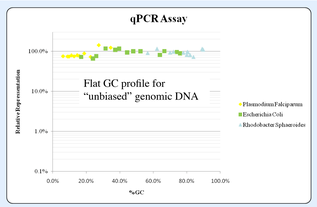
Courtesy of Gaddy Getz, Cancer Program, Broad Institute

Goals

- Test and compare bias between different thermal cyclers.
- Test if slower ramp (Δ°C/sec) speed decreases bias.
- Test the standard PCR protocol and the optimized PCR protocol.
- Test GC bias during Whole Genome Amplification (WGA)

Assay

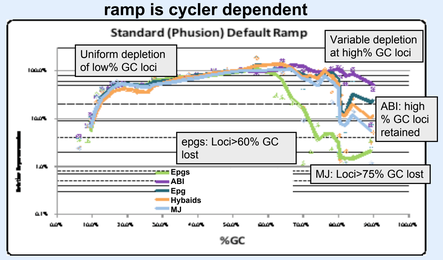
- Mix of 3 genomes as input material
  - *Plasmodium falciparum* (19% GC)
  - *Escherichia coli* (51% GC)
  - *Rhodobacter sphaeroides* (69% GC)
- Relative abundance of 36 loci from 6% to 90% GC measured by quantitative PCR (qPCR).



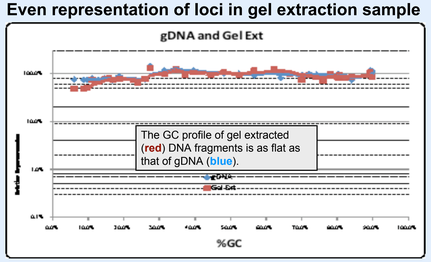
Results: Polymerase Chain Reaction (PCR)

- Bias introduced after gel extraction by subsequent PCR amplification.
- Standard PCR protocol is sensitive to the model of the PCR machine used.
- Optimized PCR protocol gives a relatively flat GC profile from ~15% to ~90% GC.

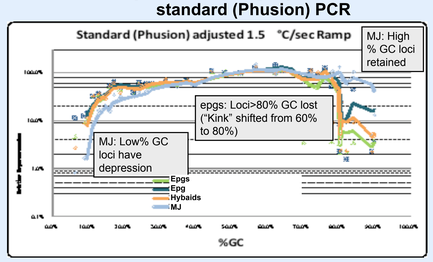
Bias during standard (Phusion) PCR at default ramp is cycler dependent



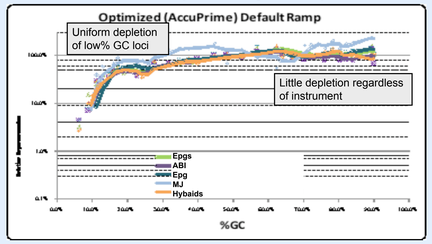
Even representation of loci in gel extraction sample



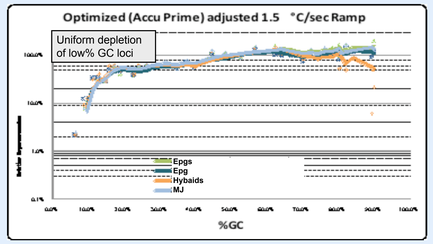
1.5°C/sec adjusted ramp slightly improves bias during standard (Phusion) PCR



Bias during optimized (AccuPrime) PCR protocol at default ramp is not cycler dependent



1.5°C/sec ramp has little affect on bias during optimized (AccuPrime) PCR

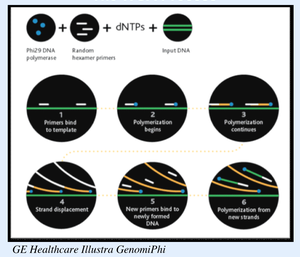


Results: Whole Genome Amplification (WGA)

The WGA Process:

- Random hexamer primers bind to denatured input DNA.
- Primers are extended by Phi29 DNA polymerase.
- Strands are displaced by proceeding primers.
- Primers bind to newly copied strands as well.

The WGA Process

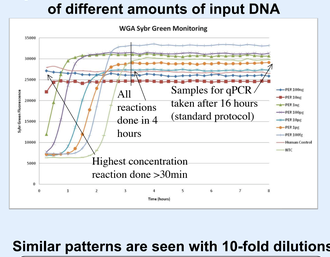


GE Healthcare Illustra GenomiPhi

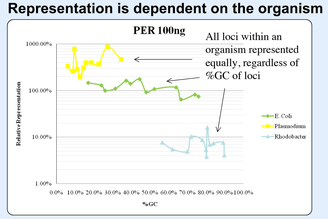
The Project Goal:

WGA is often used to amplify small quantities of genomic DNA prior to sequencing. We used the same mix of genomes and qPCR assays to study bias during WGA. Amounts of input DNA was varied from 1pg to 100ng.

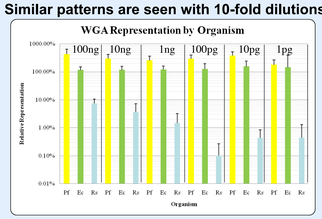
Sybr Green as a real-time detector for WGA of different amounts of input DNA



Representation is dependent on the organism

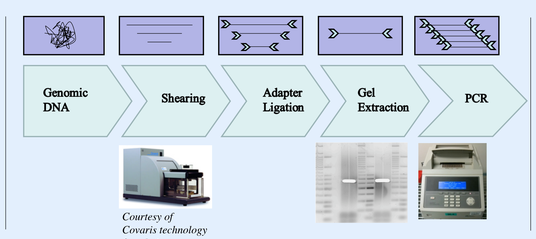


Similar patterns are seen with 10-fold dilutions



Materials and Methods

Illumina Library Preparation



Courtesy of Covaris technology foundation

PCR Protocols

Standard: Phusion DNA polymerase
1x 30"/98°C
10x 10"/98°C
30"/65°C
30"/72°C
1x 5min/72°C

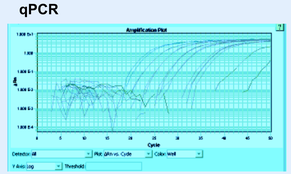
Optimized: AccuPrime Taq HF DNA polymerase
1x 3min/98°C
10x 1:20min/98°C
1:20min/65°C
1x 10min/65°C

PCR Machines

	Ep Gradient S	ABI	Ep Gradient	Hybaids	MJ
Default Ramp Speed	6°C/sec ↑ 4.5°C/sec ↓	2.2°C/sec ↑ 2.2°C/sec ↓	4°C/sec ↑ 3°C/sec ↓	3°C/sec ↑ 2°C/sec ↓	3°C/sec ↑ 3°C/sec ↓
Adjusted Ramp Speed	1.5°C/sec ↑ 1.5°C/sec ↓	Optimized	1.5°C/sec ↑ 1.5°C/sec ↓	1.5°C/sec ↑ 1.5°C/sec ↓	1.5°C/sec ↑ 1.5°C/sec ↓

Analysis of Bias via qPCR

qPCR



Conclusions

- PCR introduces bias in the Illumina Library Preparation protocol.
- Standard PCR protocol (Phusion High Fidelity with short denaturation) can have varying degrees of bias, depending on the thermocycler used. Bias on poorly performing fast-ramping instruments can be slightly improved by slowing ramping speed to 1.5°C/sec (further optimization is needed for better performances)
- Optimized PCR protocol (AccuPrime Taq High Fidelity) has very little bias regardless of thermocycler. Slowing the ramping speed to 1.5°C/sec has little to no effect on performance.
- WGA has bias too, however it appears to depend on the GC content of the organism's genome, rather than the %GC of each individual locus.

Acknowledgments

Special thanks go to the Genome Sequencing Platform for access to the machines used in this study. Thanks also go to Megan Rokop, Rachel Woodruff and Allison Martino for the input and advice that went into this poster.