Globally identifying mutations in *Plasmodium falciparum* using high-density tiling microarrays

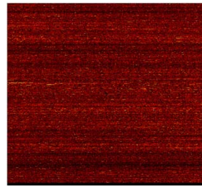
Neekesh V. Dharia, Scott Westenberger, Elizabeth A. Winzeler

The Scripps Research Institute, La Jolla, California

Overview

The implementation of suitable methods for globally identifying mutations in *Plasmodium falciparum* associated with drug resistance would aid our ability to prescribe effective therapies, develop new drugs and set public health guidelines. Recent efforts have been made to learn more about the genetic diversity of *P. falciparum*, the parasite responsible for the most severe form of human malaria. The genome of a laboratory clone, 3D7, was sequenced in 2002, and the sequencing of a handful of other strains since has shown high levels of genetic variation. However, sequencing of different strains has not yet become widespread and is a time consuming process. With our custom high-density Affymetrix microarrays that tile through the genome of the reference strain 3D7, we are able to detect single nucleotide polymorphisms (SNPs), amplification events and deletions accurately, rapidly and cost effectively for various laboratory and clinical strains of *P. falciparum*.

Microarray Design



Our custom Affymetrix array contains >4.8 million probes of approximately 25 nucleotides each tiling through the genome of 3D7 at 2 or 3 base pair spacing.

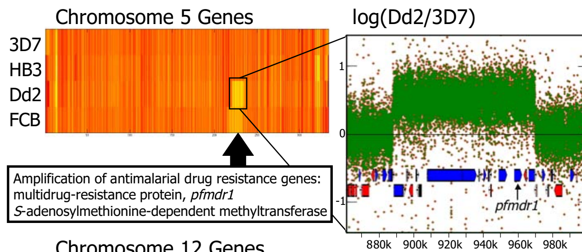
```
AAGAACCAACTTTTATCGAGCTCTT
   AACCAACTTTTATCGAGCTCTTAAT
      CAACTTTTATCGAGCTCTTAATGAA
         CTTTTATCGAGCTCTTAATGAATAC
            TTATCGAGCTCTTAATGAATACCCC
               TCGAGCTCTTAATGAATACCCCAAA
                  AGCTCTTAATGAATACCCCAAAGAT
                     TCTTAATGAATACCCCAAAGATTGT
                        TAATGAATACCCCAAAGATTGTTAC
```

Above: Hybridization of 3D7 genomic DNA.

Right: Each nucleotide position is sampled multiple times in each experiment.

Amplification Events

Method: We detect amplification and deletion events by comparing probe intensities for a gene to the mean for the genome on the same chip. For localizing the boundaries of the event, we compute the log ratio of intensities with respect to a reference 3D7 hybridization after normalization.



Chromosome 12 Genes

3D7
HB3
Dd2
FCB

Amplification of folate biosynthesis gene GTP cyclohydrolase I

deletion or polymorphic | single copy | amplification

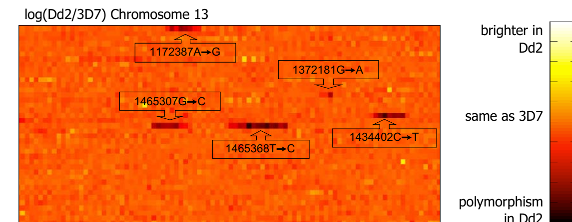
Validation: We demonstrate with high resolution amplification events previously described:

- amplification of *pfmdr1* in Dd2 and FCB
- amplification of GTP cyclohydrolase in 3D7, HB3, Dd2, FCB

Application: Quickly determine amplification and deletion events in one hybridization experiment.

SNP Detection

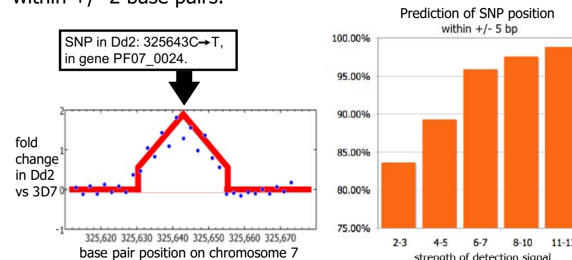
Method: The microarray data to be used for the analysis is normalized by the non-linear method with a baseline array; the invariant probes are fit to a non-linear relation defined by the running median line. Next, a paired t-test is performed using a sliding window of three overlapping probes to determine if the intensities are different in the sample compared to two reference 3D7 hybridizations. Combining the data from the sliding windows, we establish the boundaries in which the polymorphism is contained. Using a model where the fold change between the sample and reference is the greatest when the polymorphism is at the center of the oligomer, we predict the position of the polymorphism.



Validation: SNP detection in Dd2 against a set of 4134 high-quality SNPs as reported by the Broad Institute:

- **>90% detection rate**
- false positive rate of 4% for two sample hybridizations
  13% for one sample hybridization

SNPs are predicted to within +/- 5 base pairs at a rate >85%. For strong signals, SNP position is predicted at a rate >85% to within +/- 2 base pairs.



Application: Rapidly detect polymorphisms in a genome.

Conclusions

We are able to accurately and rapidly detect genomic changes, including single nucleotide polymorphisms, amplifications and deletions, in *P. falciparum* with one or two hybridization experiments.

Our method of probing genetic variability in *P. falciparum* will allow for a wider and deeper understanding of how the parasite is evolving both in the laboratory and worldwide, facilitating efforts to more effectively treat malaria.

Future directions:

- Survey of *P. falciparum* evolution worldwide and in the lab.
- Identify rates of evolution for all genes and all epitopes.
- Determine changes associated with drug resistance or other phenotypes in laboratory or clinical strains.

References:

Gresham *et al.* (2006). Science. 311(5769):1932-1936.

Kidgell *et al.* (2006). PLoS Pathogens. 2(6):562-576.

Volkmann *et al.* (2007). Nature Genetics. 39(1):113-119.