Scratching away at *Plasmodium falciparum* in Papua New Guinea.
Who, what, and where...?
Malaria

- Numerous malaria species
  - infecting a range of hosts
  - *Primates, Rodents, Birds, Lizards* …

- 5 species infect Humans
  - *P. falciparum* and *P. vivax*
    - cause majority of deaths
  - *P. malariae*, *P. ovale*
    - generally mild infection, rarely fatal
  - *P. knowlesi* – zoonotic

- Multiple cross species transfer events
Malaria

Plasmodia
- Protozoan – unicellular eukaryotic organism
- Genome size 20-30MB
  - 16 chromosomes, including ~6Kb Mitochondrion, and 30-40Kb Apicoplast
  - Exosome-like vesicles (DNA content unknown)
- Mosquito vector
  - Many vector species
- Parasite life cycle involves two hosts
- Multiple concurrent species and clone infections
- Has **haploid** and diploid phases
  - Haploid in humans
  - Diploid in mosquito
Malaria life cycle: delightfully complex

- Multiple phases and stages of infection
- Multiple concurrent clonal infections
- Genome replicates in Liver and Red blood cells
- Haploid in human host
- Fertilization of female gamete occurs in the mid gut of the mosquito (Recombination occurs)
The continuing fight against malaria

- Quinine purified from cinchona tree
- Laveran discovers malaria parasite (1907 Nobel Prize)
- Ross demonstrates mosquitoes transmit malaria
- Hans Andersag discovers CQ
- 1st global malaria eradication program
- 1st documented case of CQR
- Program abandoned
- 37 countries declared malaria free
- Erad’n program
- Faltering control
- Control
- Scale-up of control
- Renewed global focus on malaria control
- Roll Back Malaria launched
- ITN shown to reduce mortality by 20%
- ACT recommended by WHO as 1st line treatment
- Establishment of GFATM
- Renewed call for eradication
- 34 eliminating countries
- 64 controlling malaria

Timeline:
- 1820
- 1880
- 1898
- 1934
- 1955
- 1957
- 1973
- 1978
- 1992
- 1996
- 1998
- 2000
- 2007
- 2014

Deaths:
- ~3.5 m deaths/yr
- ~0.5 m deaths/yr
- 1.5-2.7 m deaths/yr
- 1 m deaths/yr
- 0.6 m deaths/yr
Malaria and Papua New Guinea

Over the last decade there has been a heightened push for control and elimination of Malaria, thanks in part to the Global Fund to combat AIDs, TB, and Malaria.

Papua New Guinea (PNG) has the highest burden of malaria outside Africa ranging from holo-endemic, to endemic to absent.

Bednets (LLIN)s and drugs (RDTs)
Study objectives

Using population genetic models and tools we aim to:

1. Map the population structure of *P. falciparum* in PNG.
   - Develop tools for molecular surveillance (Barcodes).
   - Apply these tools to understand the dynamics of the population structure
     - eg, identify “hot spots” (sources of outbreaks).

2. Investigate genetic signatures in the genome.
   - How have external forces, such as drugs and the introduction of bed nets, effected change in the genome and population structure:
     - Genetic diversity
     - Linked genetic characters
     - Selection landscape
     - Bottle necks
     - Expansions
PNG: a diversity hotspot

The diverse Geography, Biogeography, Host, Mosquito spp., and malaria control conditions throughout PNG produces diverse micro-epidemiologies.

- Historical separation & isolation of human populations
  - Genetic heritage, cultural behaviour, and movement
- Mosquito
  - 11 anophelene species
- Biogeography
  - Diverse fauna and flora (5% of worldwide diversity)
- Geography
- Control
  - Past and present
Microsatellite studies have found
- Moderate to high population structure,
  - suggesting fragmented demes
- High levels of genetic diversity across all populations
- Variable allelic (microsatellite) richness
  - lower richness in inland populations

Prevalence of infection
- No significant difference between catchment areas,
- High variation between villages ranging from 1.7% to 43.5%
- Pre bed-net distribution the prevalence was more uniform,
- Post bed-net distribution prevalence was more variable with “hot spots” interspersed with relative low prevalence.
Mapping the population structure of *Plasmodium falciparum*

These initial works are excellent and informative but…

- Microsatellites have limitations
  - Do not know inherited history
  - There are issues with quality control of data

- Whole genome sequencing (WGS) and Single Nucleotide Polymorphisms (SNPs) provide a more robust data set to draw inferences from.
  - SNPs give a more detailed picture of evolutionary relationships between individuals
  - Can identify specific SNP markers for different biological questions and applications
Plasmodium whole genome sequencing & SNPs

- MalariaGen consortium (Sanger, Hinxton, UK)
- 100s of researchers over many countries
- 1000s genomes sequenced from across the globe
- To date identified 680 000 quality SNPs in *P. falciparum*

- PNG *P. falciparum*
  - 132 whole genomes sequenced
  - ~ 90 single clone infections
  - 54 Madang sequences, collected ’08
  - 50 East Sepik sequences, collected ’13
  - 29 Milne Bay sequences, collected ’13
- **Each sample** typed for 680 000 SNPs
**WANTED: Informative SNPS**

**SNP selection**
- ~15000 variant SNPs identified in PNG
- 6025 Minor Allele Frequency (MAF) > 0.1
- 5786 with high read coverage

**Discard SNPs** under selection pressure or in linkage disequilibrium
- (BayeScan, and other $F_{ST}$ estimates, Tajima’s $D$)
- Rank SNPs by $F_{ST}$ values
- Select top 96 SNPs evenly distributed across the genome
- Select 96 neutral SNPs randomly distributed

**Validate results**
- PCA and phylogenetic analyses

**Design Fluidigm assay**

---

**PCA & Phylogenetic analyses from our 192 iSNPs**

*MAP of Papua New Guinea showing regions with SNP data.*
**Mapping P. falciparum across PNG**

LLIN and RDTs Malaria Control assessment

- 10,156 blood spots collected 2008
- 81 villages across 17 Provinces
- 3784 *P. falciparum* infections
- 2349 typed for number of clones per infection
- 585 single clone infections
- 549 dual clone infections

### Table: Malaria Control assessment sample collection sites

<table>
<thead>
<tr>
<th>Province</th>
<th>Samples screened</th>
<th>Pf positive</th>
<th>MSP2 positive</th>
<th>Single infections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bougainville</td>
<td>470</td>
<td>34</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Central</td>
<td>393</td>
<td>41</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Chimbu</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>East New Britian</td>
<td>366</td>
<td>73</td>
<td>29</td>
<td>17</td>
</tr>
<tr>
<td>East Sepik</td>
<td>844</td>
<td>329</td>
<td>260</td>
<td>76</td>
</tr>
<tr>
<td>EHP</td>
<td>446</td>
<td>73</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Gulf</td>
<td>330</td>
<td>59</td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td>Madang</td>
<td>1480</td>
<td>367</td>
<td>301</td>
<td>133</td>
</tr>
<tr>
<td>Manus</td>
<td>277</td>
<td>98</td>
<td>30</td>
<td>16</td>
</tr>
<tr>
<td>Milne Bay</td>
<td>55</td>
<td>11</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>Morobe</td>
<td>1938</td>
<td>466</td>
<td>271</td>
<td>148</td>
</tr>
<tr>
<td>New Ireland</td>
<td>411</td>
<td>157</td>
<td>76</td>
<td>43</td>
</tr>
<tr>
<td>Northern (Oro)</td>
<td>712</td>
<td>190</td>
<td>26</td>
<td>14</td>
</tr>
<tr>
<td>West New Britian</td>
<td>626</td>
<td>198</td>
<td>140</td>
<td>86</td>
</tr>
<tr>
<td>West Sepik</td>
<td>756</td>
<td>138</td>
<td>46</td>
<td>22</td>
</tr>
<tr>
<td>Western</td>
<td>616</td>
<td>51</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>WHP</td>
<td>422</td>
<td>64</td>
<td>13</td>
<td>8</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td><strong>10156</strong></td>
<td><strong>2349</strong></td>
<td><strong>1233</strong></td>
<td><strong>585</strong></td>
</tr>
</tbody>
</table>
**Mapping P. falciparum population structure across PNG**

- Fluidigm 96.96 Dynamic Array™ IFC genotyping
- 792 samples, approx. 900 clones, will each be typed for 192 iSNPs

Using our genotype barcode the plan is to map *P. falciparum* population structure and then investigate dynamics
- *eg.*, flow between mainland and different Islands areas
- *eg.*, the origins of strains involved in epidemic outbreaks in highlands areas.
- *eg.*, origins of imported or locally acquired cases among the workforce of resource companies (mining companies)
The global fight against malaria

Global mortality has been reduced by 40%
3.3 million lives saved

52 countries are on track to reduce their malaria case incidence rates by 75%

But the effect of these efforts need to be monitored and maintained...

P. vivax, P. falciparum and acquired immunity

“We shall defend our island, whatever the cost may be, we shall fight on the beaches, ... we shall never surrender.”
Churchill
Accurate knowledge and understanding of a pathogen’s epidemiology and population structure is vital in any effort to control and eventually eliminate it.
Future direction: Chicken scratching

2. Investigate genetic signatures in the genome and changes in the population dynamics.

How have external forces, such as drugs and the introduction of bed nets, effected change in the genome and population structure:

- Bottle necks
- Expansions
- Genetic diversity
- Linked genetic characters
- Selection landscape
- Where? When? Why?
**Chicken scratching's**: Investigate signatures in the genome and changes in the population dynamics.

- Population dynamics and changes
  - Measure, map, and date population changes (Skyline plots)
  - 132 WGS, 3 sites
    - 15,000 nuclear SNPs - recombining
    - 6K Mitochondrion, 30-40KB Aplicoblast – non recombining
  - ~ 900 clones, nation wide
    - 192 Nuclear SNPs

- Multitude of programs ....
  - SNAPP, BEAST
  - Not all SNPs were created equal
    - *P. falciparum* is 70-90% AT rich
  - What’s new toys are there in the play ground?
Acknowledgements

Madang team:
Livingstone Tavul,
Eline Kattenberg, Leanne Robinson,
Inoni Betuela,

Goroka team:
Peter Siba, Jo Wapling, Justin Pulford
Sarah Javati, Lilah Tol, Elma Nate, Lincoln Timinao

Alotau team:
Albina Teliki, Tulaponi Setepano, Jennifer Gogomira, Karrol Sylvester, Thomas Mark,
Gurney HC staff & CEO Billy Naidu of PHA Milne Bay Province;
Children & parents within the Gurney Catchment area,

Maprik Team:
Dorish Walsh, Nathan Yalabus, Robert Gitimbu,
Lindy Markem, Lawrence Rare

Alyssa Barry,
Barry and Mueller group
Celine Barnadas, Mita Hapsari Hazairin,
Sofonias Tessema, Connie Li Wai Suen
Anita lerch

Melanie Bahlo
Bahlo group
Natacha Tessier

Systems Biology
Liz Miller
Dominic Kwiatkowski,
Magnus Manske
Olivo Miotto
Plasmodium population structure in PNG

work flow

- Sequence whole genomes of samples collected across the north coast of Papua New Guinea (PNG)

- Identify 192 putatively biogeographically informative single nucleotide polymorphisms (SNPs)

- SNP type single clone Plasmodium infections collected from 17 provinces across PNG
  - Collected as part of Malaria control program (MalCon)

- Develop barcodes from most informative SNPs

  - Investigate underlying biological questions from genomic data and apply to National data set (MalCon)
Identifying geographically informative SNPs

WANTED SMART SNPs

From the 700k high quality bialleic SNPs we will identify and select ~200 that are geographically informative to screen our population sample set with.

- Minor allele frequency (MAF) ≥0.1 in PNG
- SNPs evenly distributed through the genome i.e., not linked or hitchhiking
- SNPs not under selection pressure or balancing pressure
  - Evolving neutrally (synonymous)
  - Tajima’s D test
- logistic regression
- High $F_{ST}$ values
- SNPs demonstrating population structure
Pf Fluidigm samples plated and pre-amped in PNG

765 selected. Predominately MOI=1 but some MOI=2
Using our Genotype barcode we will investigate and characterize

- Samples collected at healthcare facilities in locations where imported or epidemic malaria is suspected

- Origin of imported or locally acquired cases among the workforce of resource companies
  - 200 odd samples from Exon

- Gene-flow between mainland and different Islands areas

- The origins of strains involved in epidemic outbreaks in highlands areas.
Depending on who you’re talking to merge the intro of Pf pop structure in PNG into one slide with multiple references and “pop outs” to describe each paper. It’s not about “numbers” but it is about what their story tells.

NEED TO INCLUDE BACK GROUND OF POPULATION STRUCTURE OF other MALARIA IN PNG NOT JUST PF AND POP STRUCTURE OF PF ELSE WHERE TO GIVE BACKGROUND ABOUT POSSIBLE STRUCTURE

AND THE MANY PURPOSES OF WHY WE DO THIS – DRUG STUDY EVALUATION, POP STRUCTURE, EVOLUTION, SELECTION PRESSURE, SELECTIVE SWEEPS,

Other examples of pop structure – drug resistance marker and that 24 SNP barcode
ii. to develop tools for molecular surveillance

**BedNet screening 2008-9**

Bill and Melinda, LLIN and RDTs  
Malaria Control assessment

- 17 Provinces, 81 villages, 13,000 samples  
~ 2500 P. falciparum infections  
~ 1000 single (monoclonal) infections

![Map of PNG with collection sites](attachment:image.png)  
**Genotyping (SNPs)**
To apply these tools to identify sources of outbreaks

- iSNPs we will genotype a nationwide data set of *P. falciparum* samples and map the population structure of *P. falciparum* across PNG.

  - The national geographic informative SNPs will be identified and a barcode produced.

- Our Genotype barcode will have the power to investigate and characterize:
  - Suspected imported malaria cases
  - *P. falciparum* migration and gene flow
  - The origins of strains involved in epidemic outbreaks in highlands areas.

The long term goal being to be able to genotype isolates from a few select SNPs and map outbreaks and local epidemics.
Malaria: Global burden

- Burden estimates
  - 219 million clinical cases
    - (CI 154 million to 289 million)
  - 660,000 deaths
    - (CI 490,000 to 836,000).

  (accurate estimates of prevalence and deaths are difficult to establish due to limitations in reporting.)

- Has been eliminated from:
  - Australia, Europe, North America - well mostly
  - DDT elimination attempts in the 1960s in developing countries not successful

- Over the last decade there has been a heightened push for control and elimination
  - bed nets, new drugs and vaccines
Malaria and PNG

Papua New Guinea (PNG) has the highest burden of malaria outside Africa ranging from holo-endemic, to endemic, to absent.
Geographically informative SNPs

AFRICA
- Kenya
- Senegal

AMERICAS
- Brazil

ASIA
- India
- Thailand
- Cambodia
- Vietnam

PACIFIC
- PNG
- Solomon Is

Global
- G
- T
- G

Regional
- G
- G
- G

Local
- A
- C
- G

Development of geographic “barcodes”