Think Big!
The bright future of linguistics

Red-crested Cardinal
Paroaria coronata

Biology today

Encyclopedia of life

Tree of Life

Tree of Life
Genbank taxonomy browser

Genbank - cytochrome b sequence

Linguistics today

Think big – scaling up...
1. Big data
2. Big methods
3. Big questions
4. Big teams

Max Planck Institute for the Science of Human History

GlottoBank
GlottoBank

LexiBank | GramBank | PhonoBank | ParaBank

D-PLACE
a global database of cultural variation linked to language trees and ecological data

Following the Comrie model

The data deluge requires computational tools

Think big – scaling up...

1. Data
2. Methods
What value can computational methods add?

- Dating language divergences
- Phylogeography
- Functional dependencies
- Networks

“linguists don’t do dates”
April & Robert McMahon (2006)

Austronesian Basic Vocabulary Database
http://language.psy.auckland.ac.nz
1201 Languages, 210 Words, 237,921 entries

Cognacy

Bob Blust  John Lynch  Jeff Marck  Malcolm Ross  Laurent Sagart

Cognate coding

<table>
<thead>
<tr>
<th>Language</th>
<th>“father”</th>
<th>cognacy</th>
<th>binary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paiwan</td>
<td>tjama</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Itbayaten</td>
<td>qamaq</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Mangarrai</td>
<td>ema</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Motu</td>
<td>tama-na</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Fijian (Bau)</td>
<td>tama-na</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Tongan</td>
<td>tama i</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Rarotongan</td>
<td>metua</td>
<td>2</td>
<td>01</td>
</tr>
<tr>
<td>Maori</td>
<td>matua</td>
<td>2</td>
<td>01</td>
</tr>
</tbody>
</table>
**Data**

- 400 well-attested languages
  - No creoles, obvious borrowing removed
- Outgroup
  - Old Chinese (controversial)
  - Buyang (less controversial)
- Binary Coding
  - Presence/absence of cognates
  - 34,440 cognate sets
  - Covariation model

**Bayesian Phylogenetic Inference**

1. Data
2. Model (and priors)
3. Tree search
4. Dating (without a strict clock)

**Uncertainty in tree estimation**

**Austronesian phylogram**
3 predictions:

Sequence, timing, pulses and pauses

Think big – scaling up...
1. Data
2. Methods
3. Questions

What are the Hilbert problems in linguistics?

David Hilbert
Martin Hiapard

Explain this!

http://www.worldmapper.org/display_languages.php?selected=583

Some suggestions...
1. Why are there approximately 7000 languages?
2. Why is language diversity distributed so patchily?
3. What drives the evolution of linguistic disparity?
4. When did spoken language evolve?
5. How far back can we push the time barrier for detecting language relationships?

Lineage through time plots

David N. Reznick & Robert E. Ricklefs,
Nature 457, 837-842 (12 February, 2009)
Do rainfall and group size drive the diversity and distribution of Australian languages?

Think big – scaling up...

1. Data
2. Methods
3. Questions
4. Teams
Big interdisciplinary teams

Vanuatu – the Galapagos of language evolution

Why do “Remote Melanesians” not look like Polynesians?
Blust (2005, 2008)

1. Phenotypic differences in “Remote Melanesia”
2. Cultural similarities
3. Language typology - serial verb constructions
4. Loss of decimal system - switch to various quinary systems
5. Large amount of sound change

Circumstantial evidence for extended contact
Distribution of retention rates

<table>
<thead>
<tr>
<th>Number</th>
<th>Paiwan</th>
<th>Cebuano</th>
<th>Maori</th>
<th>Nengone</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ita</td>
<td>usa</td>
<td>tahi</td>
<td>sa</td>
</tr>
<tr>
<td>2</td>
<td>dusa</td>
<td>duhá</td>
<td>rua</td>
<td>rewè</td>
</tr>
<tr>
<td>3</td>
<td>tjelu</td>
<td>tulo</td>
<td>toru</td>
<td>tini</td>
</tr>
<tr>
<td>4</td>
<td>sepatj</td>
<td>upát</td>
<td>whaa</td>
<td>ece</td>
</tr>
<tr>
<td>5</td>
<td>rimá?</td>
<td>lima</td>
<td>rima</td>
<td>xduŋ</td>
</tr>
</tbody>
</table>

Vanuatu PNG highlands Chimbu Valley, New Guinea Mek warrior, Irian Jaya Pentecost Island, Vanuatu Tanna, Vanuatu
Vanuatu – the Galapagos of language evolution