

A more detailed early Austronesian phylogeny

Laurent Sagart

CNRS, Paris

BACKGROUND

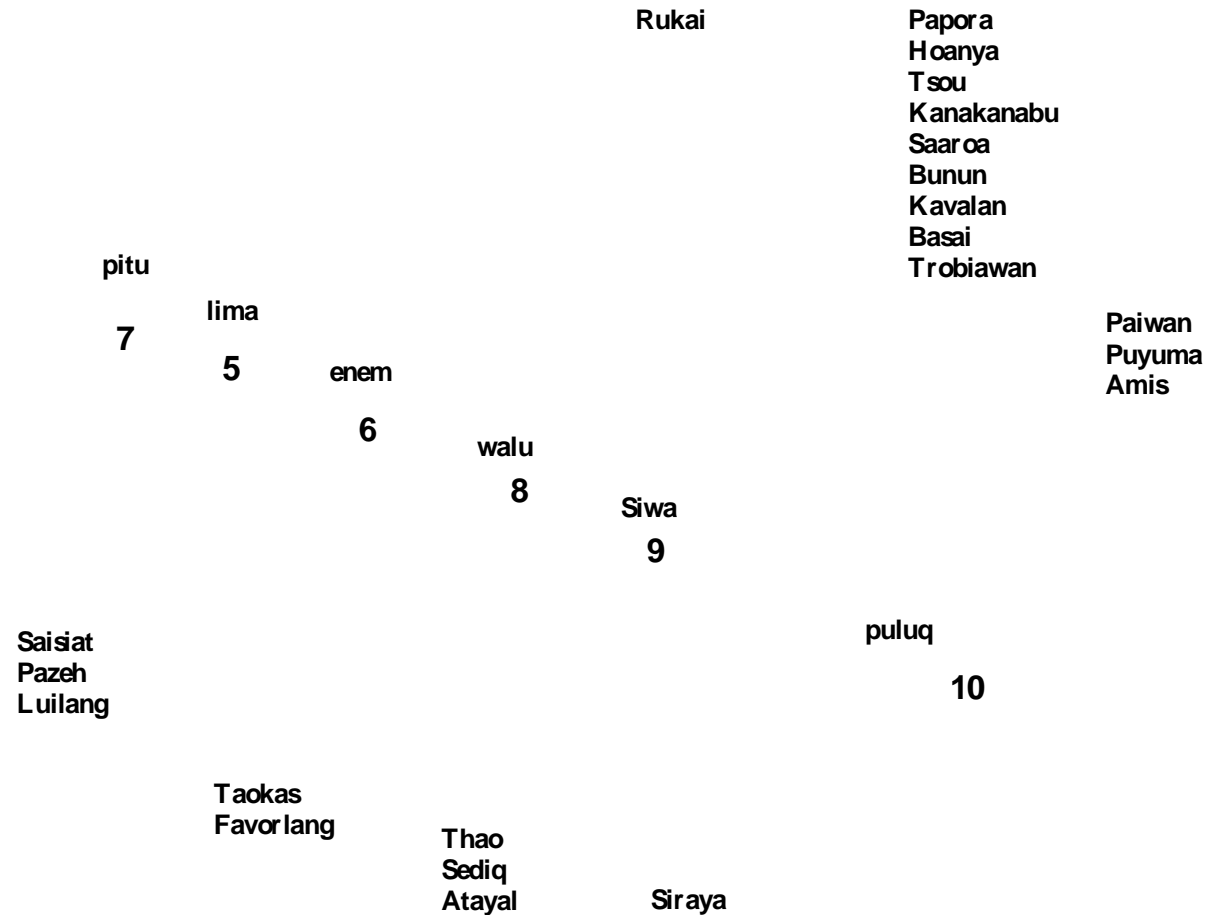
Compatibility methods (Meacham and Estabrook (1985))

- Look for mutually compatible characters
- Mutually compatible=compatible with the same tree
- The largest set of mutually compatible characters defines the best tree
- Mutual compatibility in itself argues that the characters are innovative
- But if the characters can independently be shown to be innovations, that strengthens the argument

PAN numerals in Blust's reconstruction

*isa '1'	UBIQUITOUS IN TAIWAN
*duSa '2'	
*telu '3'	
*Sepat '4'	
*lima '5'	SPATIALLY NESTED IN TAIWAN
*enem '6'	
*pitu '7'	
*walu '8'	
*Siwa '9'	
*puluq '10'	

Spatial nesting of isoglosses in Taiwan

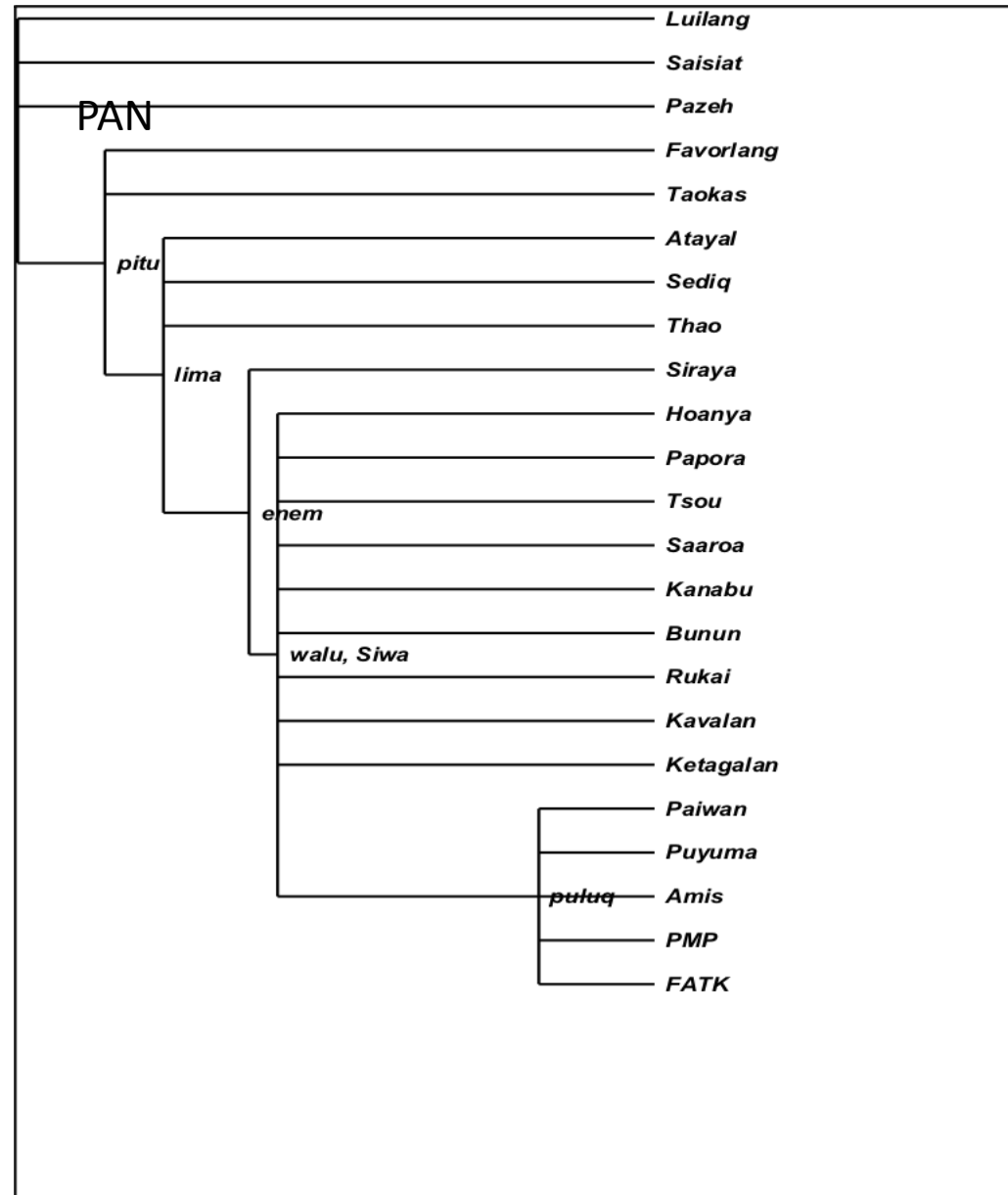


By definition, nested characters are mutually compatible

- for 6 adjacent characters in a single subsystem to be nested is not likely the product of chance
- The following tree expresses the numeral-based phylogeny

Sagart, L. (2008) the expansion of setaria farmers in east asia: a linguistic and archaeological model

- 6 innovations
- one innovation per node (in one case, two)
- **Today I will add 33 innovations that are compatible with this tree**



Evidence that 3 of the 6 numerals appeared after PAN

I presented etymologies for

- *pitu '7' < *RaCep-i-tuSa '5 + 2'
- *walu '8' < *RaCep-a-telu '5 + 3'
- *Siwa '9' < *RaCep-i-Sepat '5 + 4'

The long forms occur in the NW language Pazehe, one of the three languages without *pitu, *walu and *Siwa

Criticism of this model

- **Winter, B.** (2010) Squib: A note on the higher phylogeny of Austronesian. *Oceanic Linguistics* 49, 1:282-287.
- **Ross, M.** (2012) In defense of Nuclear Austronesian (and against Tsouic). *Language and Linguistics* 13, 6: 1253-1330.
- **Blust, Robert A.** 2014. The higher phylogeny of Austronesian and the position of Tai-Kadai: another look. The 14th International Symposium on Chinese Languages and Linguistics (IsCLL-14), Taipei, 2014.



nesting pattern is a coincidence

This presentation

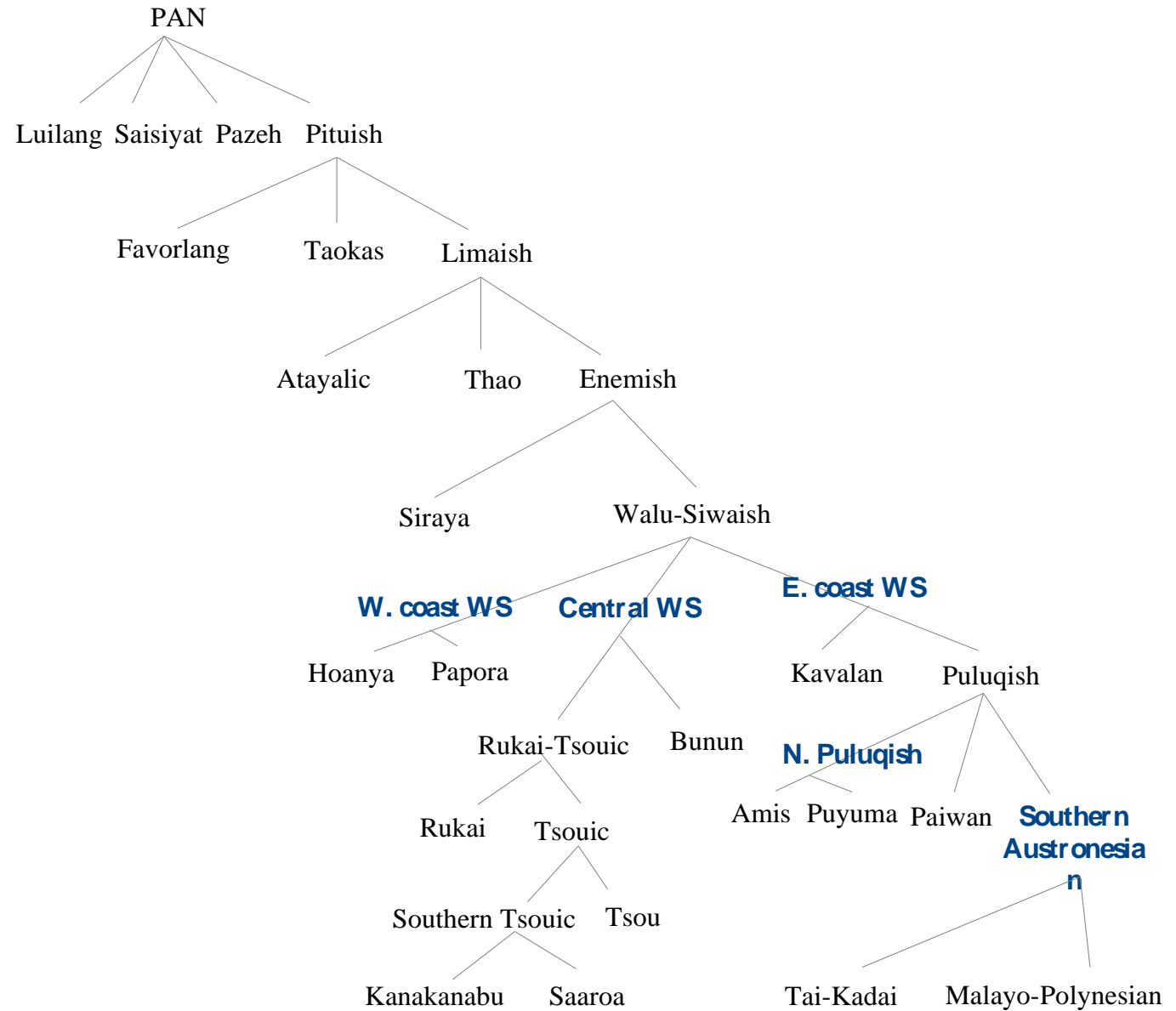
- 39 mutually compatible characters
- ability to increase number of compatible characters shows nesting pattern not a coincidence

In addition

Etymologies will be presented for *enem '6' and *sa-puluq '10'

A more detailed phylogeny

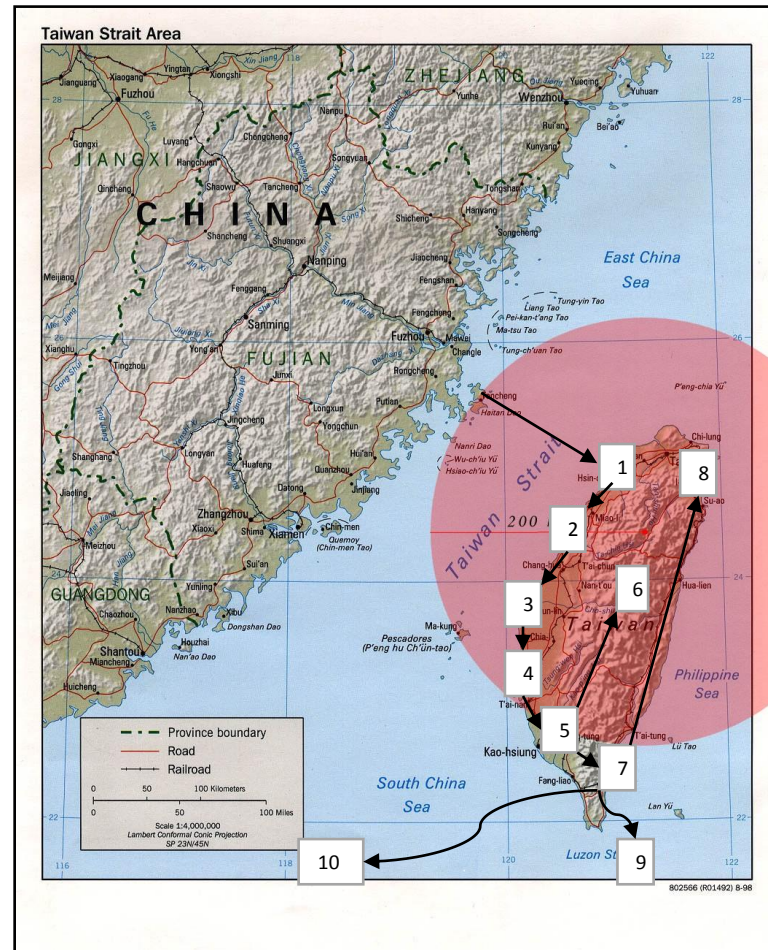
5 new nodes



number of innovations, by type

lexical		
	numerals	14
	basic vocabulary	14
	cultural vocabulary	4
	subtotal	32
phonological		
	merger	1
	metathesis	1
	vowel syncope	1
	analogical change	1
	subtotal	4
morphological		
	gain of a verbal marker	1
	gain of numeral markers	2
	subtotal	3
	total	39

The spatial pattern of Austronesian settlement of Taiwan



An online resource

- open-access
- evolutive: permitting corrections and additions
- able to receive feedback from users
- allowing non-linear exploration via hyperlinks

The solution: blog posts

- The evidence for my tree is distributed over some twenty posts on my blog 'Sino-Tibetan-Austronesian'
- Forming a network connected by hyperlinks

To access the resource, go to the URL in the abstract to this presentation:

https://f-origin.hypotheses.org/wp-content/blogs.dir/3902/files/2020/12/Clickable-Austronesian-phylogeny_2020-.pdf

or write to me asking for the URL:

laurent.sagart@gmail.com

Highlights

Etymology of *puluq ‘10’

- Modern Amis has a bound morpheme *polo* ‘to separate from and leave alone’ (Namoh 2013)
- Homophonous with *polo* ‘ten’, both < *puluq

Etymology of *puluq ‘10’

- In counting above 10 in a decimal system, one isolates n sets of ten units and ***mentally leaves them alone*** before expressing the remainder
- for instance ‘35’ can be expressed as ‘3 sets of ten, leave them alone, and 5’
- *puluq ‘to leave alone’ is the source of *puluq ‘10’

from *baCaqan to *puluq

- The word for '10' in Proto-Eastern-Walu-Siwaish was *baCaqan
- '35' in that language may have been **telu-baCaqan-puluq-lima** '3 times 10, leave those alone, five'.
- *baCaqan was redundant, so one could say simply **telu-puluq-lima** '35'
- with *puluq acquiring the meaning 'ten'

Etymology of *enem '6'

- Here is part of Blust's ACD entry for 'PAN' *enem '6':

PAN	*enem	six
Formosan		
Papora	ne-nom	six
Tsou	nomə	six
Kanakanabu	u-nə'mə	six
Saaroa	ənəmə	six
Proto-Rukai	*eneme	six
Puyuma (Tamalakaw)	enem	six
Paiwan	enem	six
WMP		
Isneg	annám	six
Itawis	annám	six
Bontok	ʔəném	six
Kankanaev	eném	six

Unexplained: Formosan forms with / or *r* instead of *n*

- Tsou #38a *lomu* '6'
- Saaroa #47a *rumu*, *urumu*, #48 *urumu* '6'
- Rukai #63 *ulum* '6'
- Paiwan #59a, 62a *urum* '6'
- Kavalan #108, 108 *arum* '6'
- Basay #106-107-108 *arum*, *rarum* '6'
- Papora #129b *rum* '6'
- Hoanya #138 *alim*, *ilim*, #142b.2 *rom* '6'

(Source: Li & Toyoshima 2006)

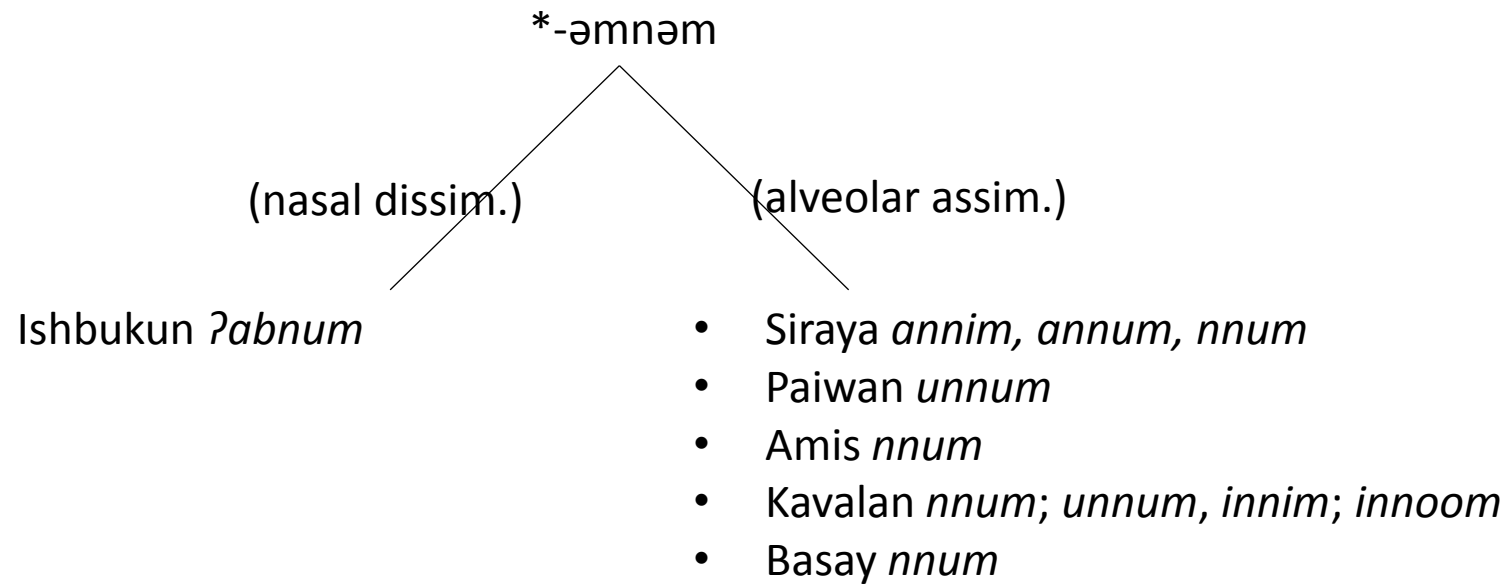
These forms reflect *-l- or *-N-, not *-n-.

Unexplained: forms with geminate -nn-

- Siraya #143 *annim, annum, nnum, ninnum, ninnim* '6'
- Paiwan #71, 72 *unnum* '6'
- Amis #82c1, 82d, 83cd, 83k.2 *nnum* '6'
- Kavalan #88b, 92a, 92b *nnum*; #92c *unnum, innim*; #101 *innoom* '6'
- Basay #102b *nnum* '6'

Remarkably, only -n- is ever geminated: there are no forms with -ll- or -rr-.

curious: Ishbukun *Bunun ʔabnum* ‘6’



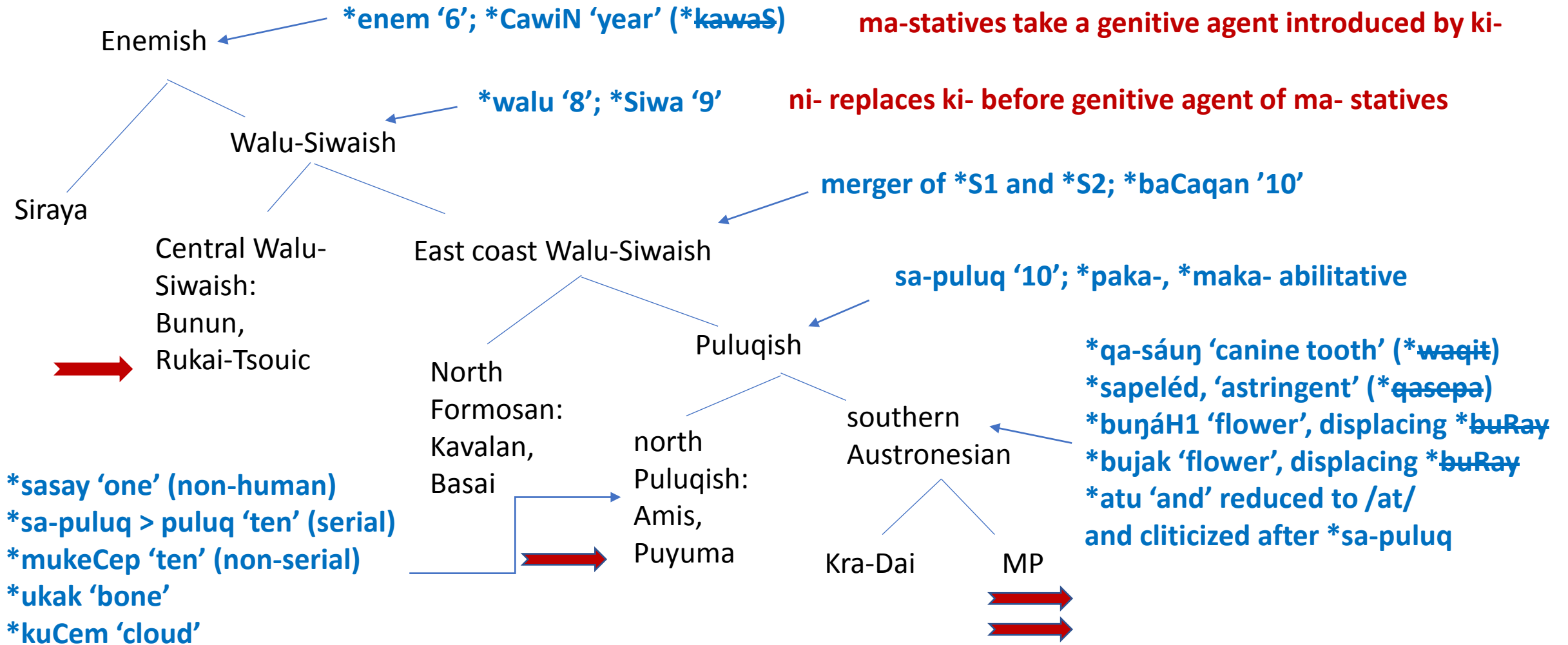
A minor root for '3': *-Nem

- Basai #112, 113 (Li and Toyoshima 2006) *pinum, peinum* (Ino) '3'
- Siraya (Li 1993) *kulom* '3', *kulom-ta?* '4'
- Makatao M3 (Tsuchida and Yamada 1991) *ra-rum-a* '3'
- Makatao M7 (Tsuchida and Yamada 1991) *nunta* '4'
- Makatao M8 (Tsuchida and Yamada 1991) *lum-ta* '4'

*(xxx)-Nem-Nem '3+3'

- Truncation: *emNem
- Nasal assimilation: *emnem '6' (-mN- > -mn-)
- Dental assimilation: *ennem
- Degemination: ***enem** '6'

A fragment of the phylogeny: from Enemish to MP



an apparent strong point of the proposal of Chen et al.

- transitivity of ma- statives defines the same collection of languages as the merger of *C and *t
- I think this is a coincidence
- unconditioned mergers spread easily to neighboring languages
- not reliable subgrouping evidence
- regularities are easy to grasp
- Unconditioned mergers in a higher language pass easily into a lower language, through bilinguals
 - merger of *n and *l from Cantonese into Hongkong Hakka
 - merger of *t^h and *h before high vowels from eastern Gan into western Min

Blust's East Formosan

- The Formosan members of the Coastal branch in Chen et al.'s proposal are Blust's East Formosan subgroup
- said to show merger of PAN *j, supposedly a palatalized velar stop [gʲ], with PAN *n
- I believe the direction of change is misconstrued
- see my presentation at 13ICAL (Taipei, 2015):

[https://www.academia.edu/14329210/East Formosan and the PAN palatals](https://www.academia.edu/14329210/East_Formosan_and_the_PAN_palatals)

The phoneme known as PAN *j was a palatal nasal n_ɸ-

- development involves fortition of n_ɸ- and denasalization
- fortition of palatal nasals has parallels in
 - Chinese (the so-called ri 日 initial)
 - Sardinian (Logudorese, Campidanese): Lat. *vinea* 'vine' > *vindza*

conclusions

- MP is not a primary branch of AN
- KD is a sister of MP

Thank you for your attention

laurent.sagart@gmail.com