

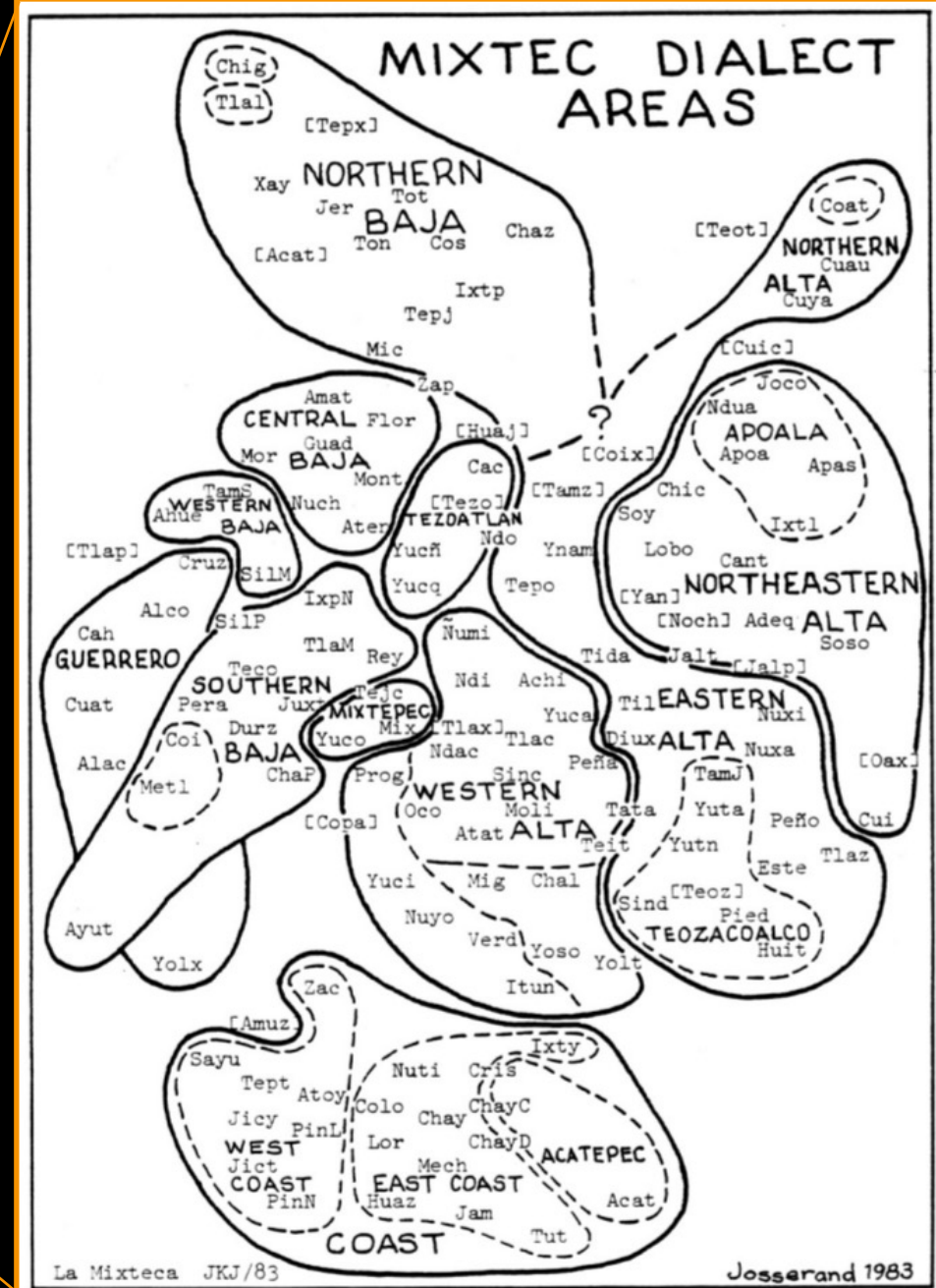
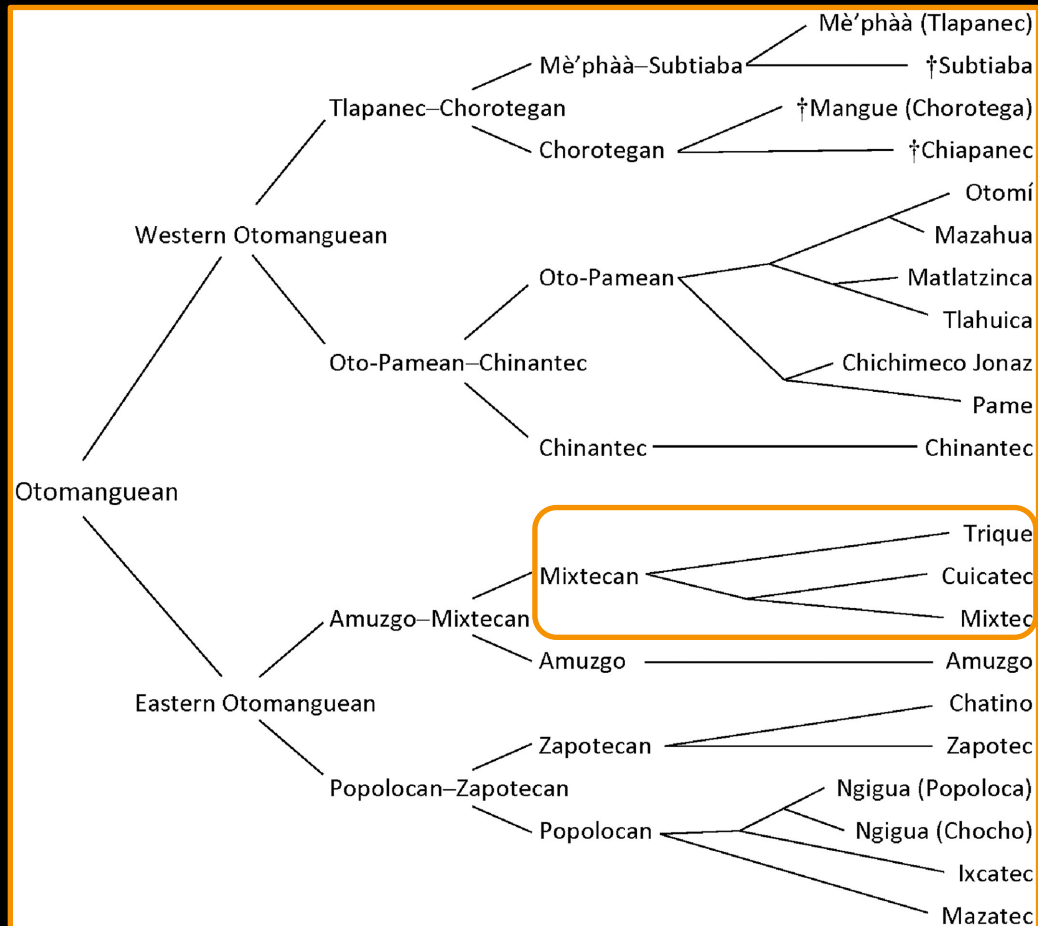
La estructura interna de la rama Mixteca: un estudio filogenético bayesiano

Sandra Auderset
Universidad de Berna

16.04.2024

Seminario permanente de lenguas mixtecas y amuzgas
Instituto de Investigaciones Antropológicas
Universidad Nacional Autónoma de México

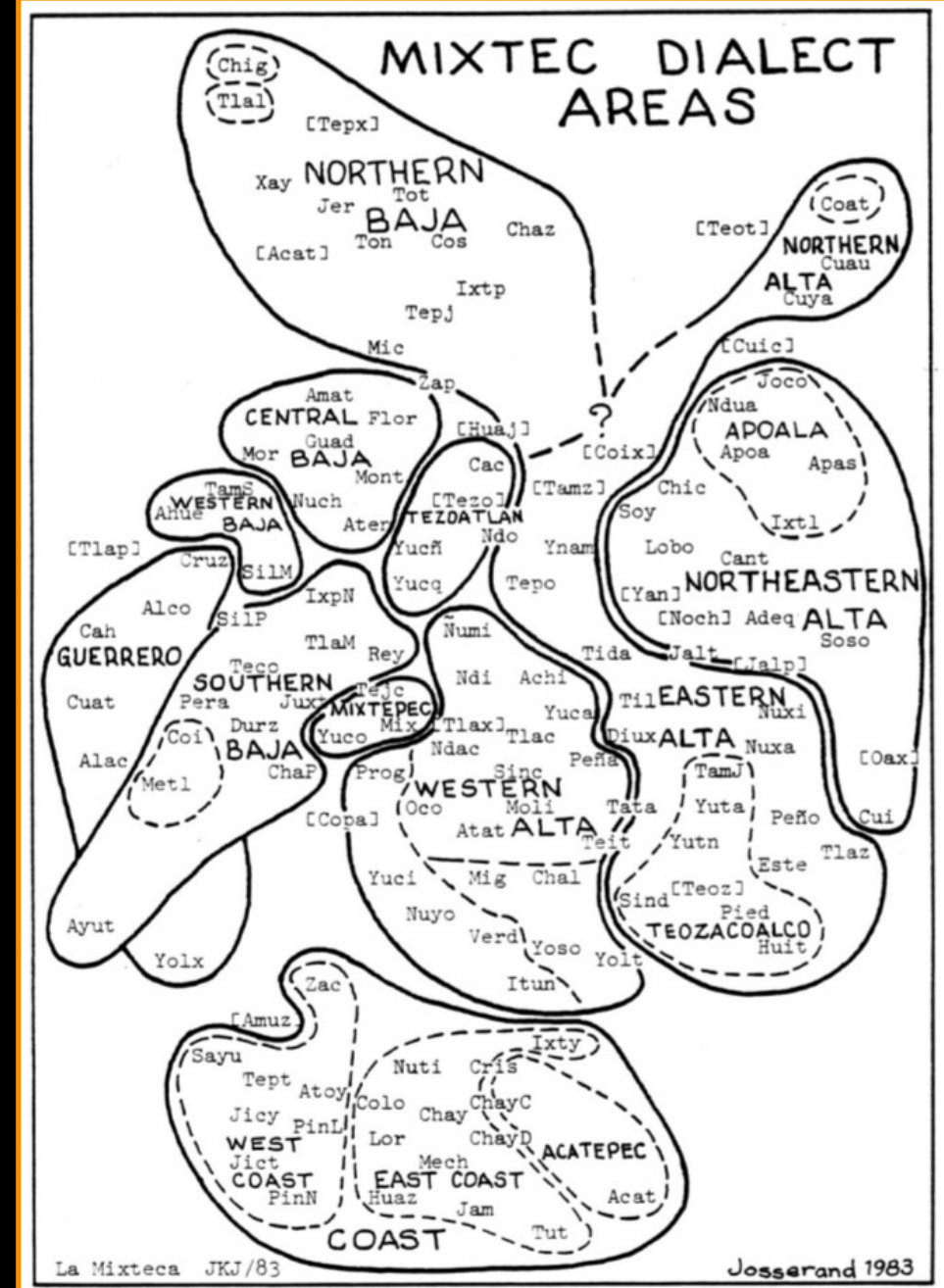
La clasificación del mixtec(an)o



La clasificación del mixteco: Áreas dialectales

Josserand 1983:

- ❖ 12 áreas, la mayoría con subdivisiones
- ❖ clasificación basada en isoglosas de correspondencias vocálicas
- ❖ datos recopilados a través de una encuesta lingüística
- ❖ 120 pueblos = 120 variedades
- ❖ 188 juegos de cognados



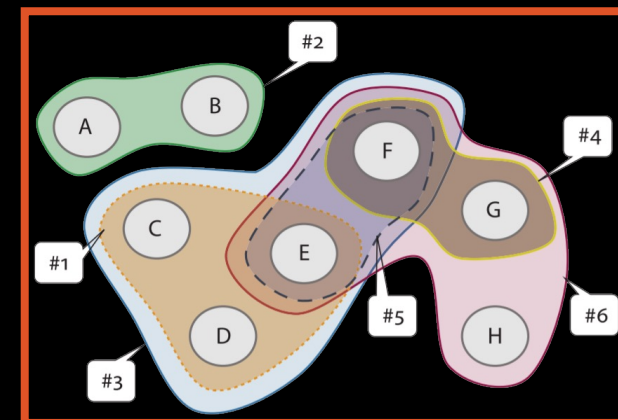
“El problema es la aparente violación o infracción de las expectativas ideales del método comparativo en la lingüística histórica (...). Pero este modelo presupone una efectiva separación de los grupos, después de un cambio que los divide. En contraste, en la Mixteca (...) vemos una situación más dinámica en el desarrollo de las familias lingüísticas.” Bradley & Josserand 1982:303

¿Árbol o onda? ¿Familia o continuo dialectal?

método comparativo:

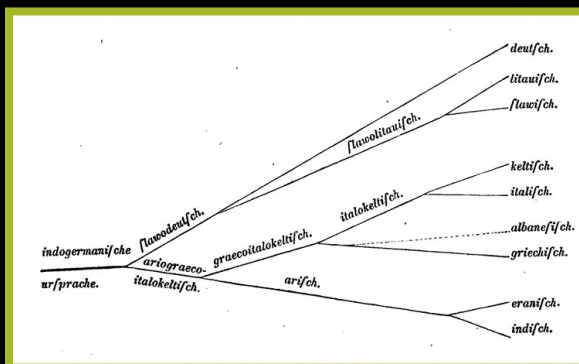
- cambio de sonido regular
- relaciones entre lenguas y subgrupos
- herencia vs. contacto
- patrones de desarrollo cómo árbol

¿desarrollo del Mixtec(an)o?



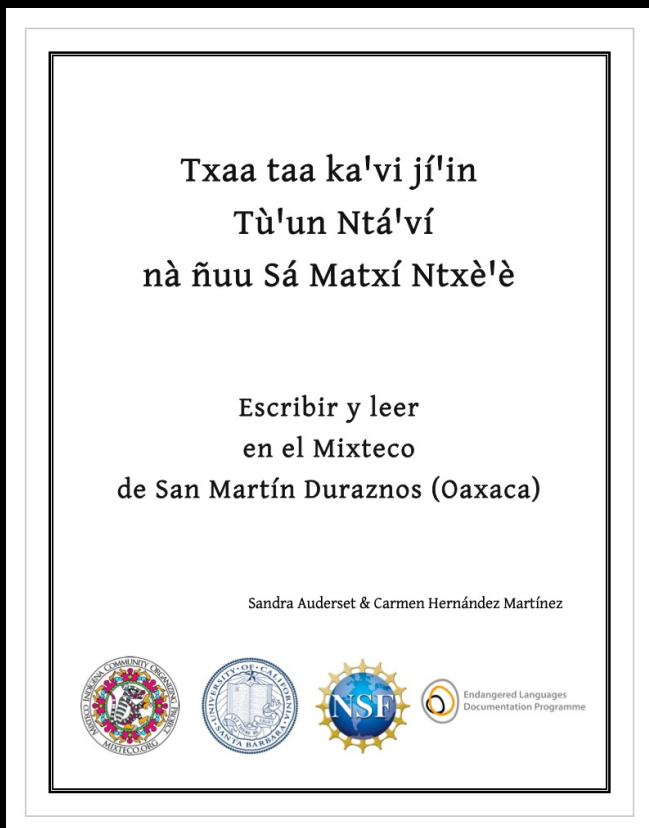
continuo/área/cadena dialectal:

- patrones de desarrollo cómo ondas, no árboles
- contacto entre lenguas relacionadas
- desafíos de la aplicación del método comparativo (Ross & Durie 1996, Kalyan & François 2018)



¿Por qué necesitamos una clasificación actualizada?

- ❖ recursos para las comunidades
- ❖ historia cultural
- ❖ arqueología
- ❖ antropología



fuentes de imágenes :

<https://www.khanacademy.org/humanities/art-americas/early-cultures/mixtec/a/mixtec-codex-zouche-nuttall>

https://commons.wikimedia.org/wiki/File:Mixtec_Carved_Stone.jpg

Metas del estudio

- ❖ recopilar nuevos datos disponibles
- ❖ combinar la lingüística histórica con la filogenética bayesiana
- ❖ punto de partida para futuras investigaciones
- ❖ identificar procesos de desarrollo (cómo árbol/cómo onda)
- ❖ reevaluación de las áreas dialectales de Josserand

❖ <https://academic.oup.com/jole/article/8/1/33/7188947>

JOURNAL ARTICLE

Subgrouping in a 'dialect continuum': A Bayesian phylogenetic analysis of the Mixtecan language family

Sandra Auderset , Simon J Greenhill, Christian T DiCanio, Eric W Campbell

Journal of Language Evolution, Volume 8, Issue 1, January 2023, Pages 33–63, <https://doi.org/10.1093/jole/lzad004>

Published: 02 June 2023

Fiologenética bayesiana aplicada a datos lingüísticos

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$

- ❖ uso del teorema de Bayes para la inferencia probabilística de árboles filogenéticos y parámetros del modelo
- ❖ subagrupamiento con probabilidad posterior (evaluación de la probabilidad de cada agrupación)
- ❖ suposiciones explícitas y modelos reproducibles
- ❖ input: juegos de cognados establecidos mediante la aplicación del método comparativo

Recopilación y análisis de los datos



lista de vocabulario básico de 209 conceptos adaptados al área cultural mesoamericana, excluyendo verbos



de fuentes publicadas (diccionarios, gramáticas, esbozos gramaticales) y fuentes inéditas (proyectos de documentación en curso)



110 variedades con listas rellenas en diferentes grados (> 18.000 entradas en total)

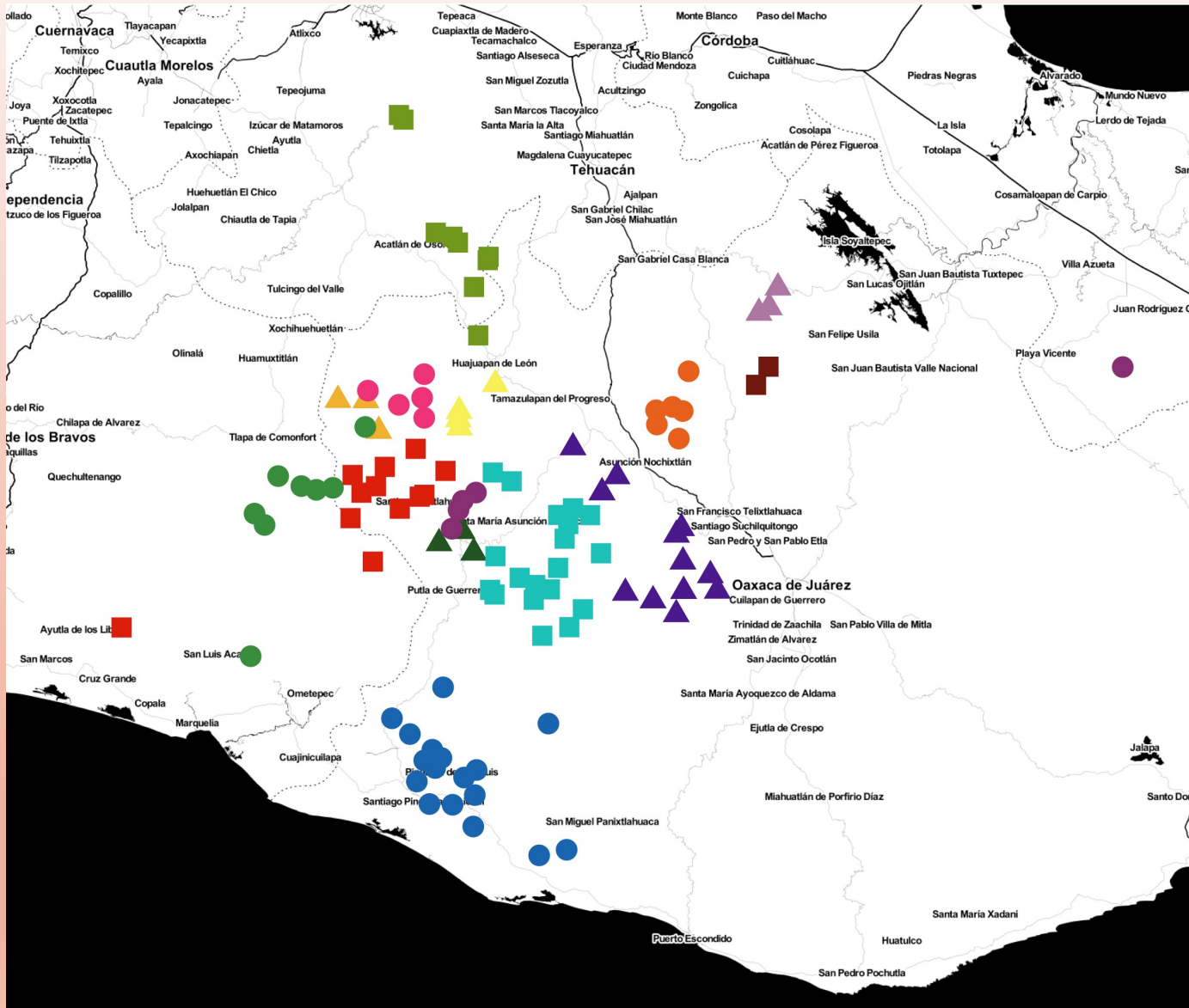


anotado para cognados parciales, tomando en cuenta el tono



clasificación manual de cognados con en el método comparativo

Muestra



Anotación de cognados

- ❖ amplio vs. fino
- ❖ matriz de cognados binarizada

Variedad	5	125	644	878
A	1	1	0	?
B	1	1	1	0
C	0	?	0	1
D	1	1	1	0

PÁJARO

VARIETY	FORM	BROAD	FINE
AbasoloValleMixtec	s a ³ a ³	49	49
AlacatlalzalaMixtec	s a ³ a ³	49	49
ChalcatongoHidalgoMixtec	s a ³ a ¹	49	49
ElJicaralMixtec	s a ³ a ³	49	49
GuadalupeVillahermosaMixtec	l a a	49	49L
IxpantepecNievesMixtec	l a a	49	49L
SanMartinDuraznosMixtec	l a ³ a ³	49	49L
SantiagoApoalaMixtec	t i + ð a a	9 49	9 49
SantiagoCacaloxtpecMixtec	t i ¹ + l a ³ a ¹	9 49	9 49L
SanMiguelElGrandeMixtec	t i ³ + s a ³ a ¹	9 49	9 49

PUS

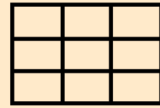
VARIETY	FORM	BROAD	FINE
CoicoyanlasFloresMixtec	ⁿ d a ³ k ^w a ¹	574	574
ElJicaralMixtec	ⁿ d a ³ k ^w a ⁵	574	574
MagdalenaPenascoMixtec	ⁿ d a ³ k ^w a ³	574	574
PiedraAzulMixtec	ⁿ d a ³ h k ^w a ¹	574	574
AlcozaucaGuerreroMixtec	l a k ^w a	574	574L
SanMiguelElGrandeMixtec	l a ³ k ^w a ³	574	574L
SanCristobalJamiltepecMixtec	l a k ^w a	574	574L
SanMartinDuraznosMixtec	l a ³ k ^w ã ¹	574	574L



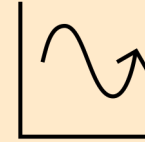
modelo de reloj
(estricto/relajado)



modelo de árbol
(BDSKY-S)



modelo de evolución de cognados
(binary covarion)



parámetros de
variación de la tasa



puntos de
calibración

MODELO



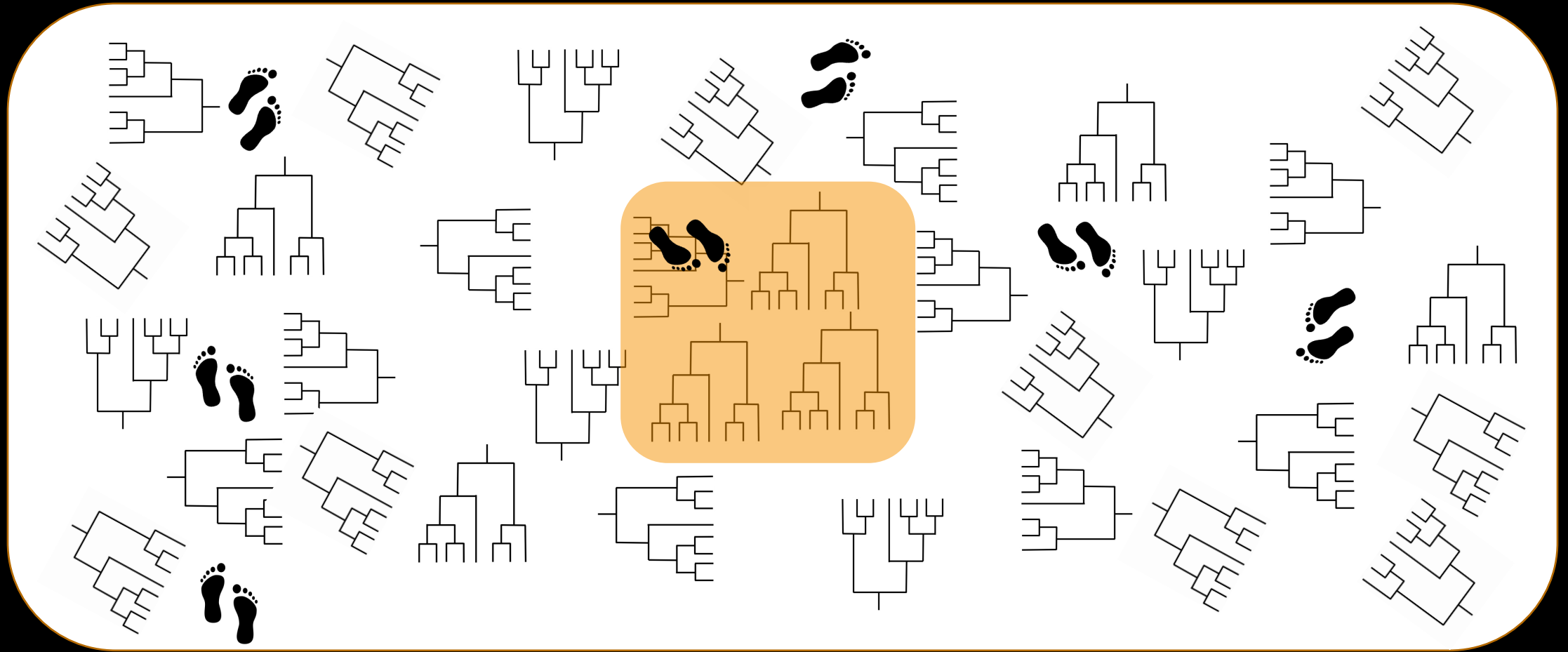
Beast2

Bayesian evolutionary analysis by sampling trees

**puntuación de verosimilitud
(likelihood score)**

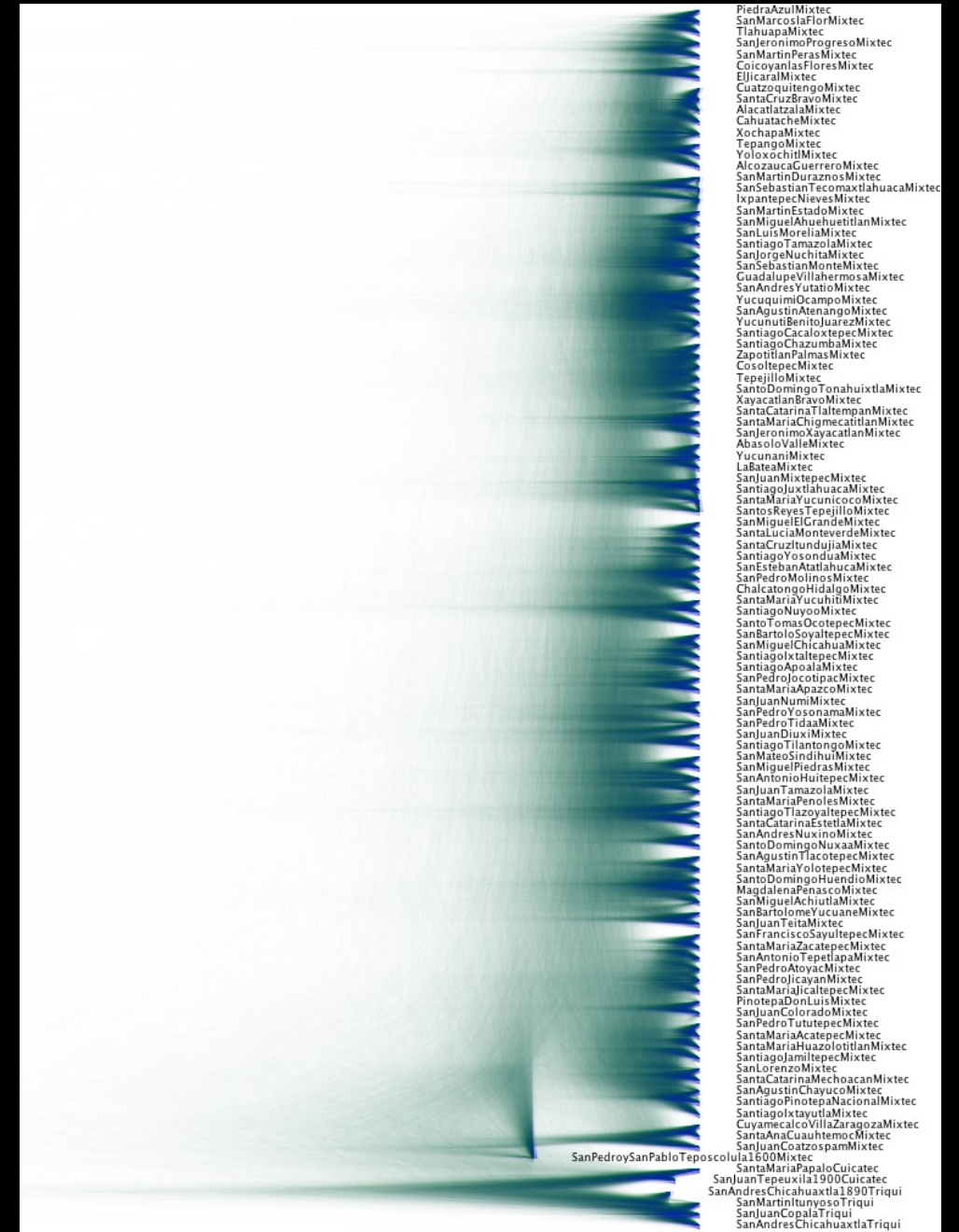
M(arkov)C(hain)M(onte)C(arlo)

Montecarlo vía cadenas de Markov

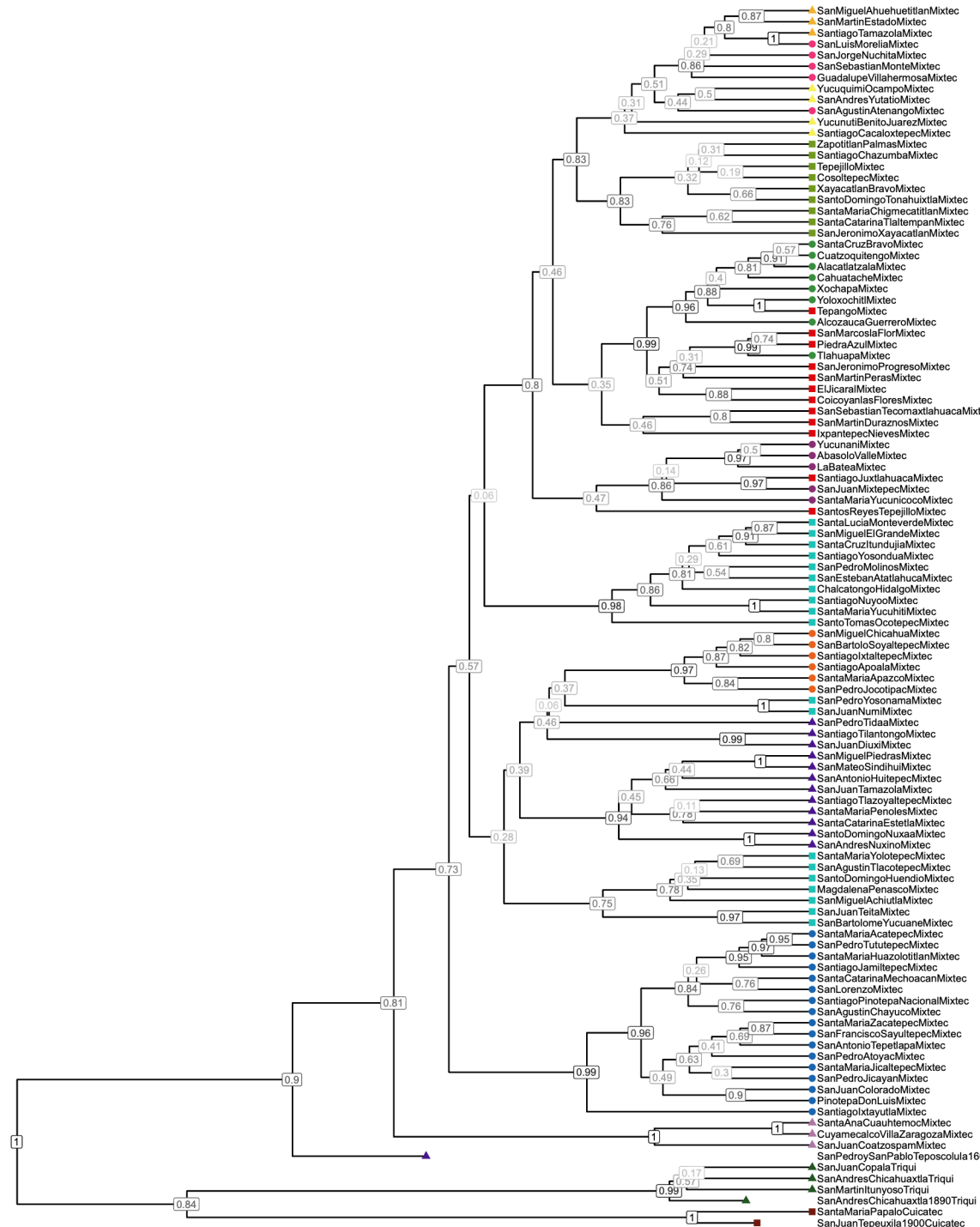


Distribución posterior de árboles

- ❖ no un árbol, sino una muestra
- ❖ muestra señal contradictoria e incertidumbre

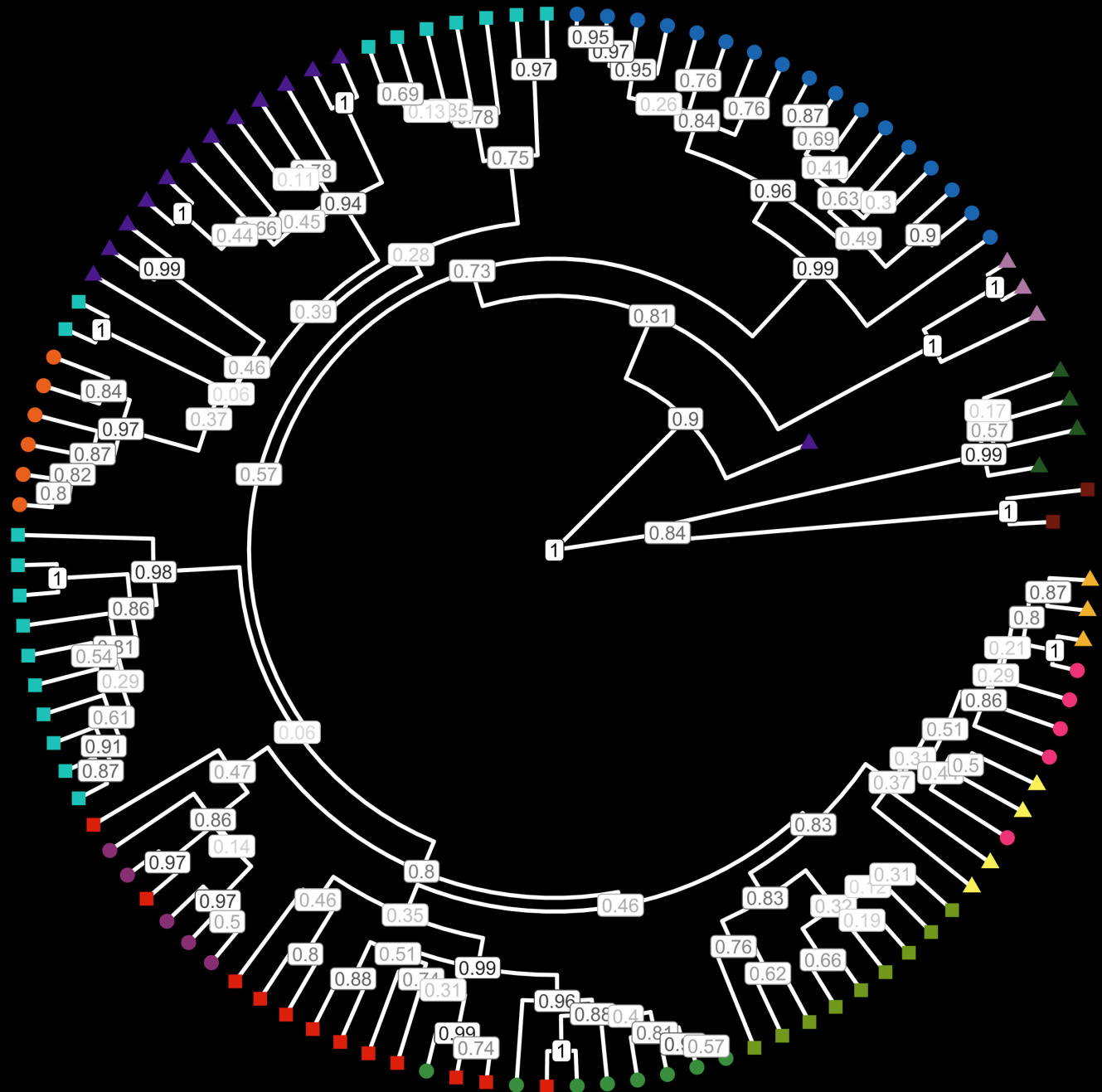


Árbol MCC

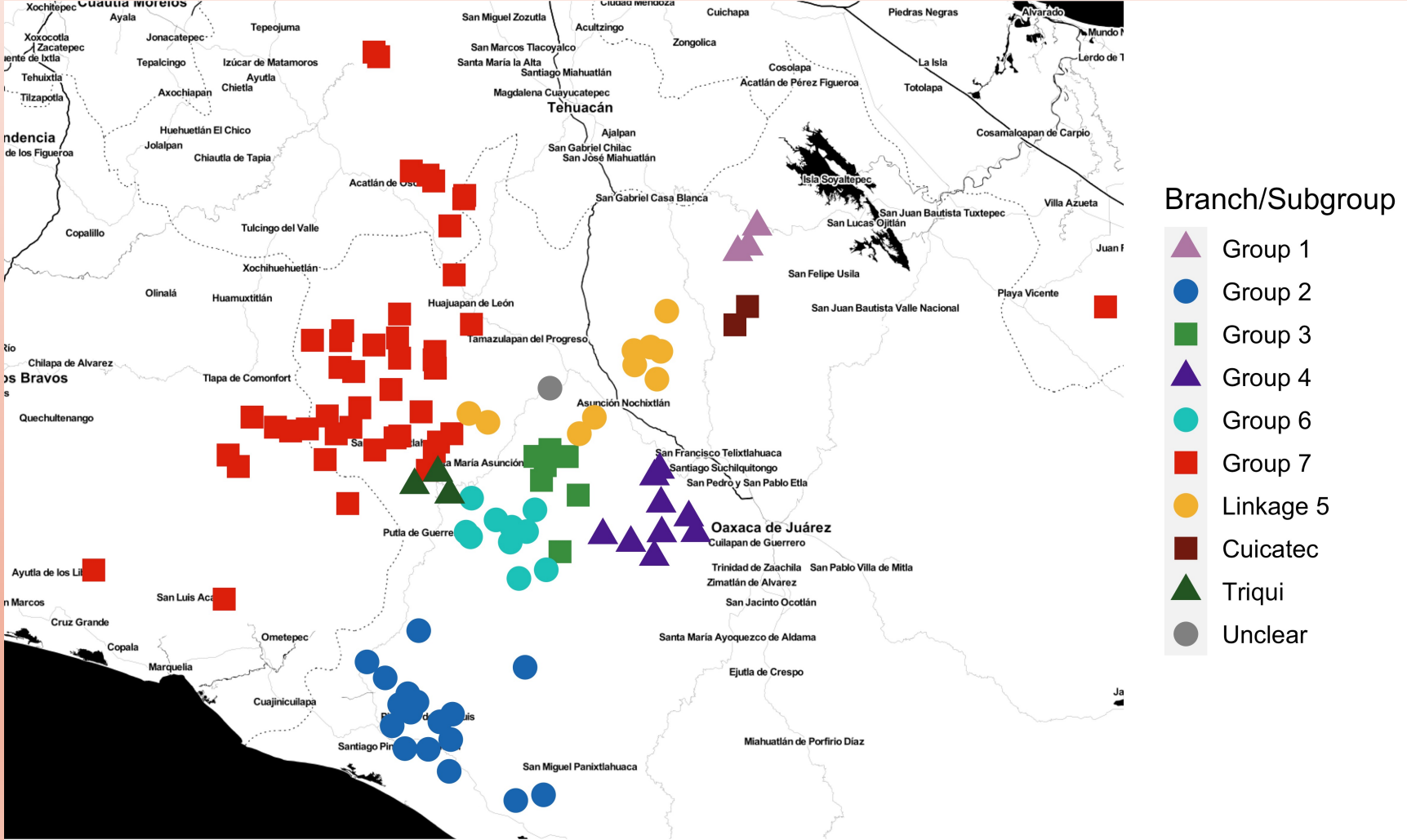


- ❖ Maximum Clade Credibility (máxima credibilidad de clados)
- ❖ muestra de árboles que mejor representan los datos dado el modelo
- ❖ probabilidad posterior para todos los nodos

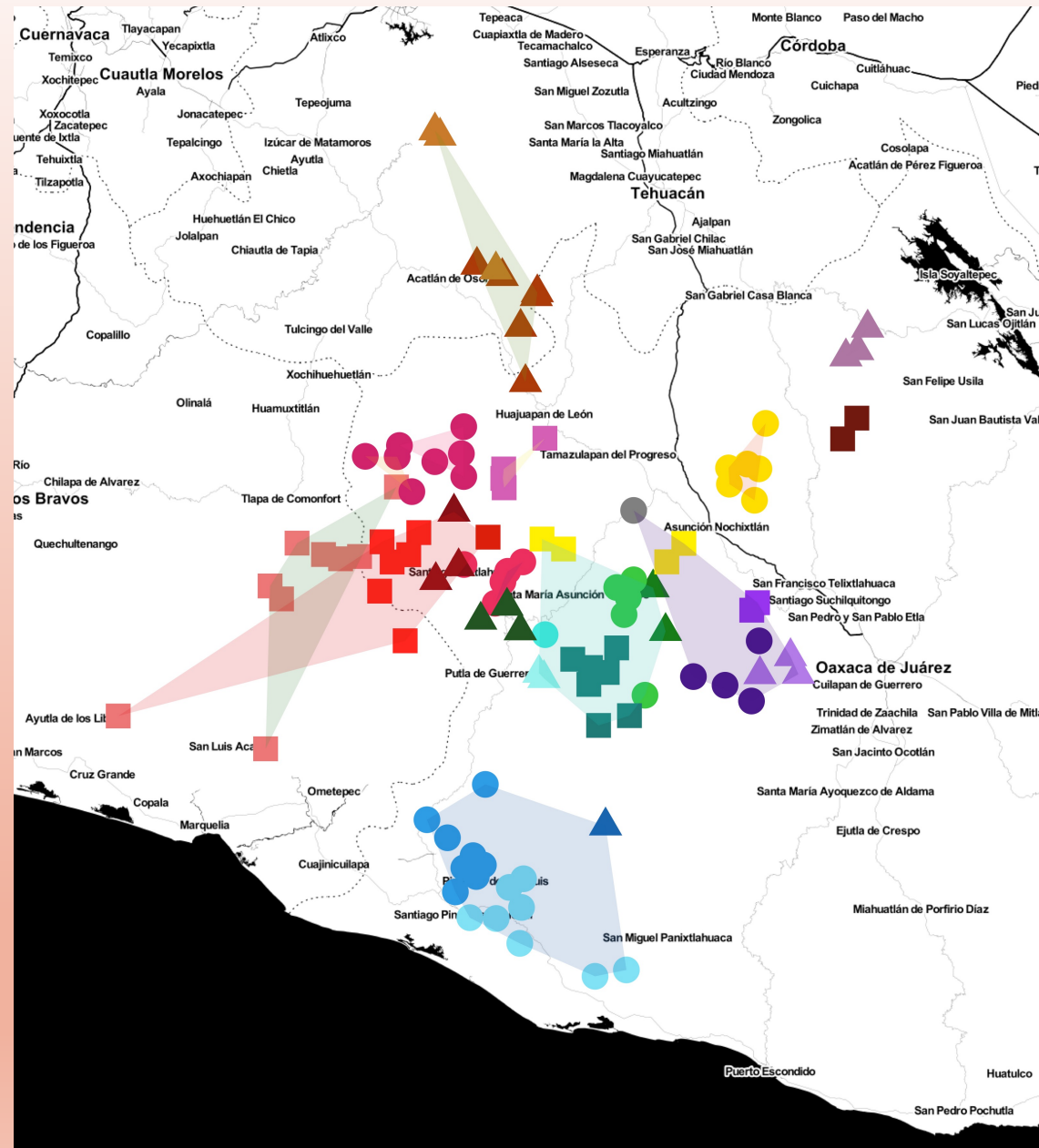
Árbol MCC



Grupos de nivel superior



Grupos nuevos vs. áreas dialectales

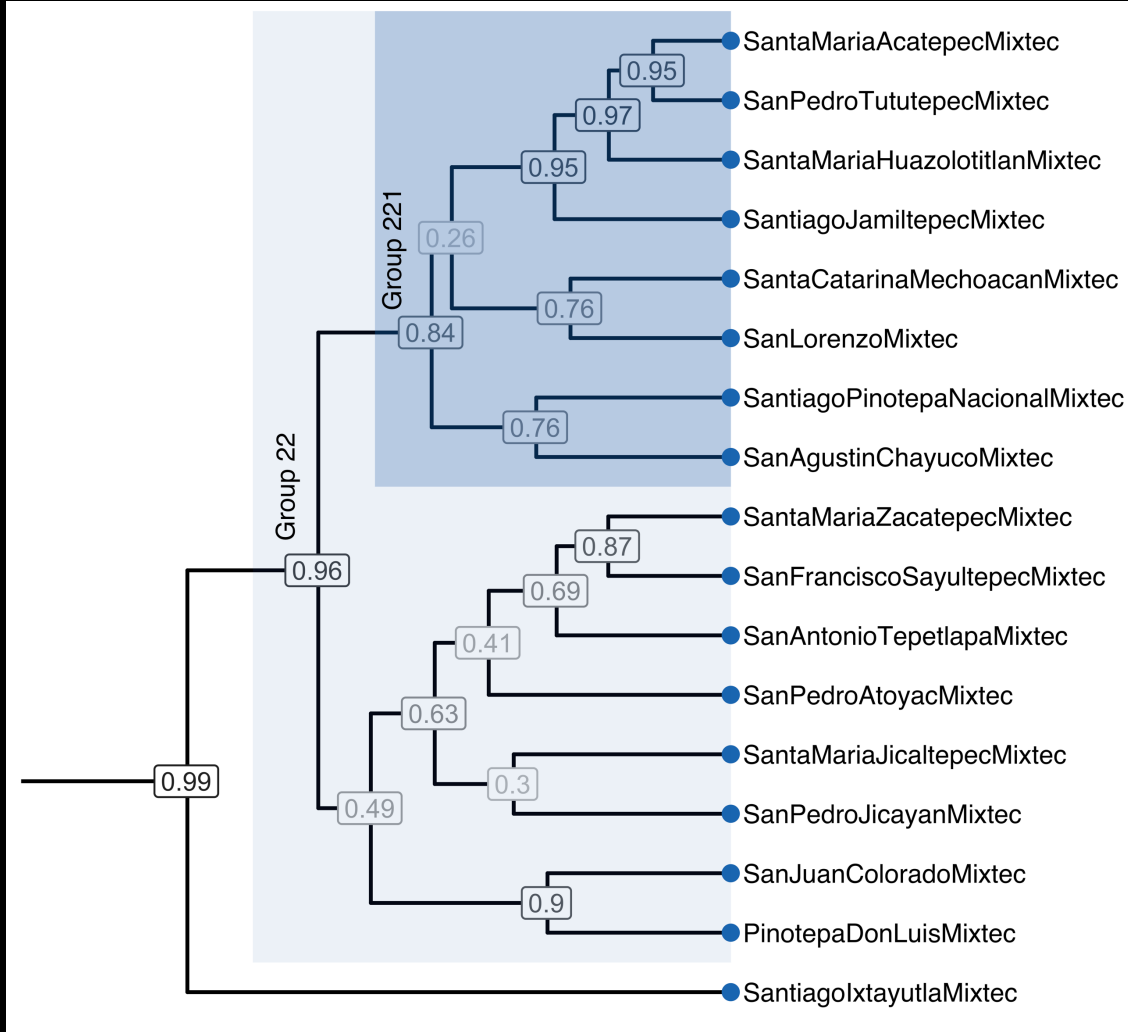


Branch/Subgroup

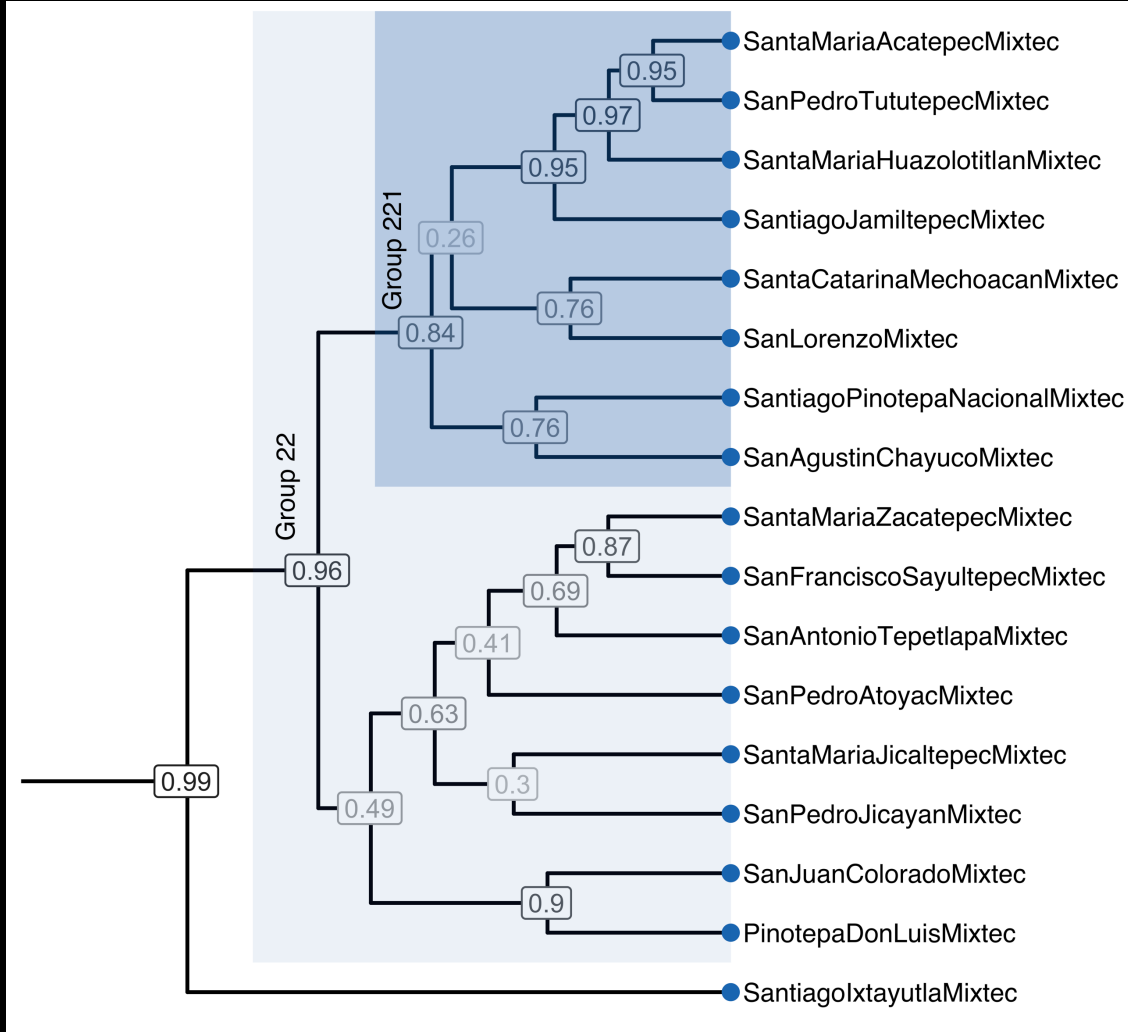
Group 1	Group 72
Group 21	Group 73
Group 22	Group 731
Group 221	Group 74
Group 31	Group 741
Group 32	Group 76
Group 41	Linkage 75
Group 42	Group 7
Group 4	Group 51
Group 61	Linkage 5
Group 621	Cuicatec
Group 622	Triqui
Group 71	Unclear

polígonos coloreados =
áreas dialectales de
Josserand

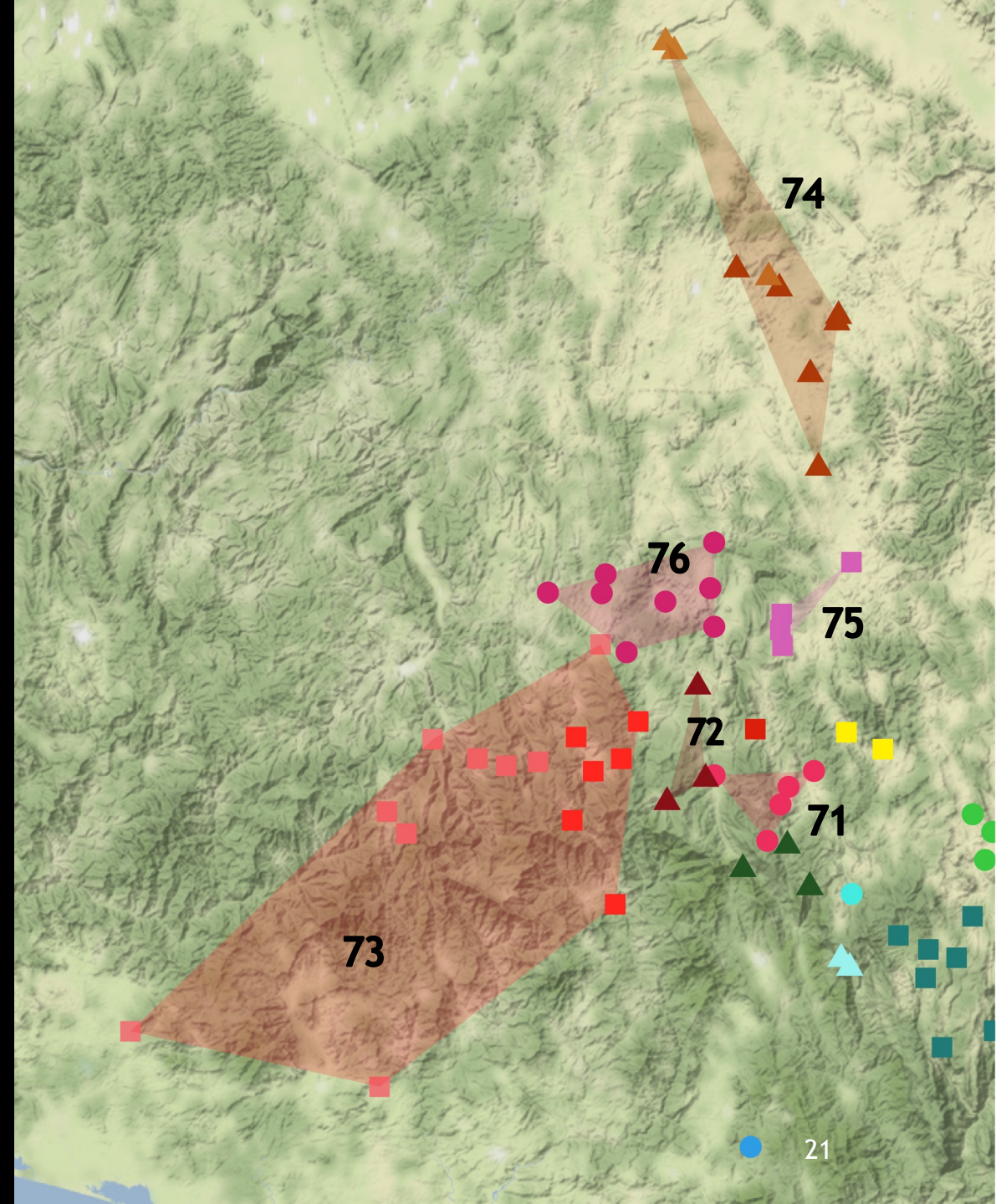
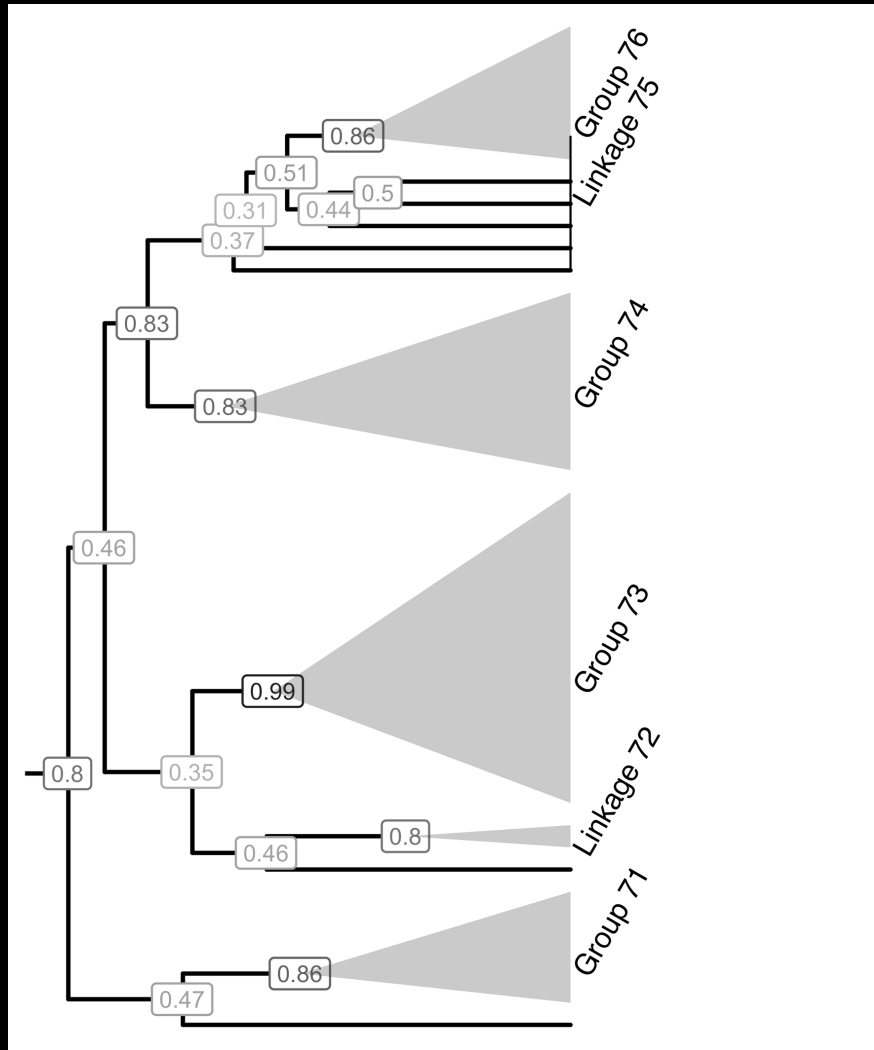
Grupo 2



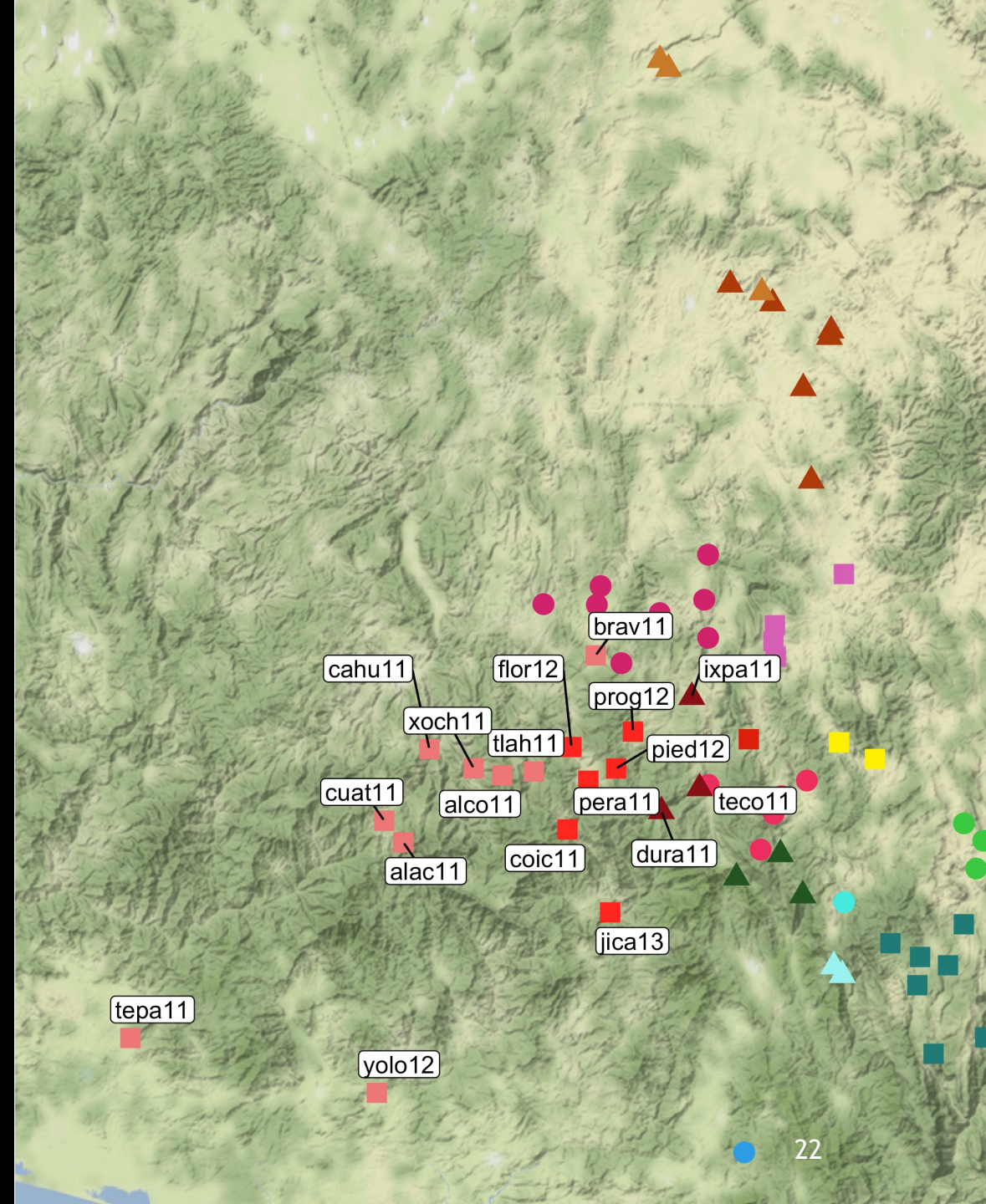
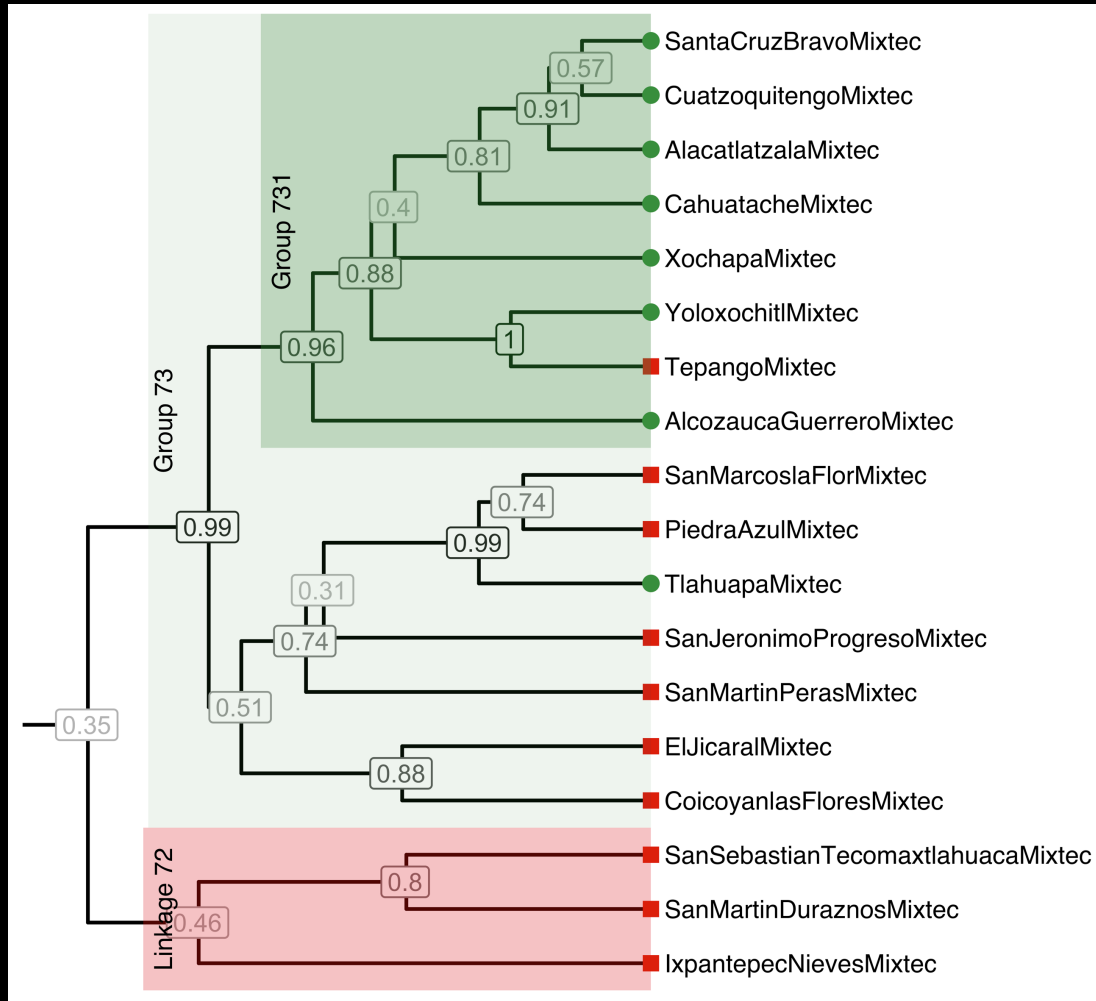
Grupo 2



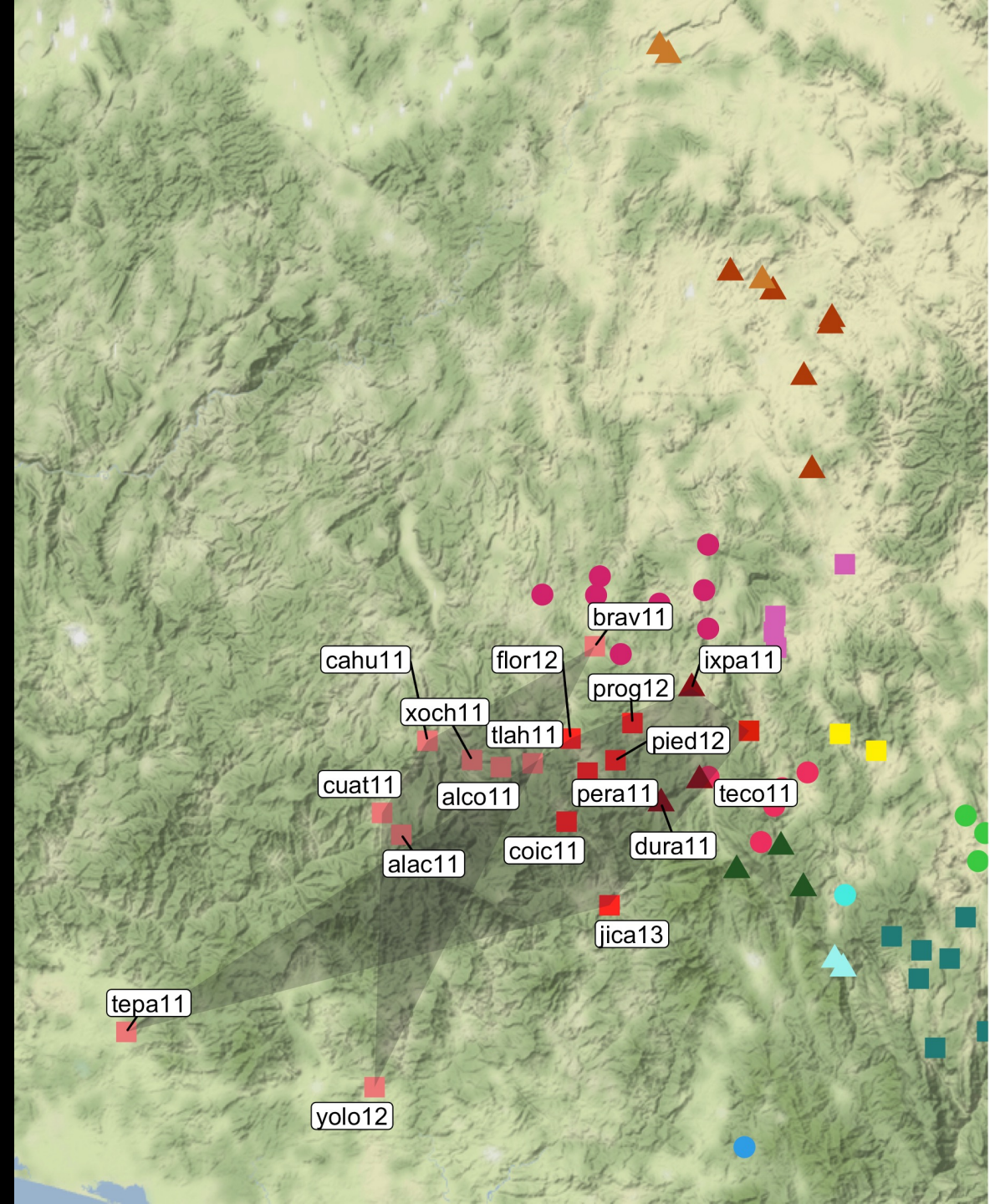
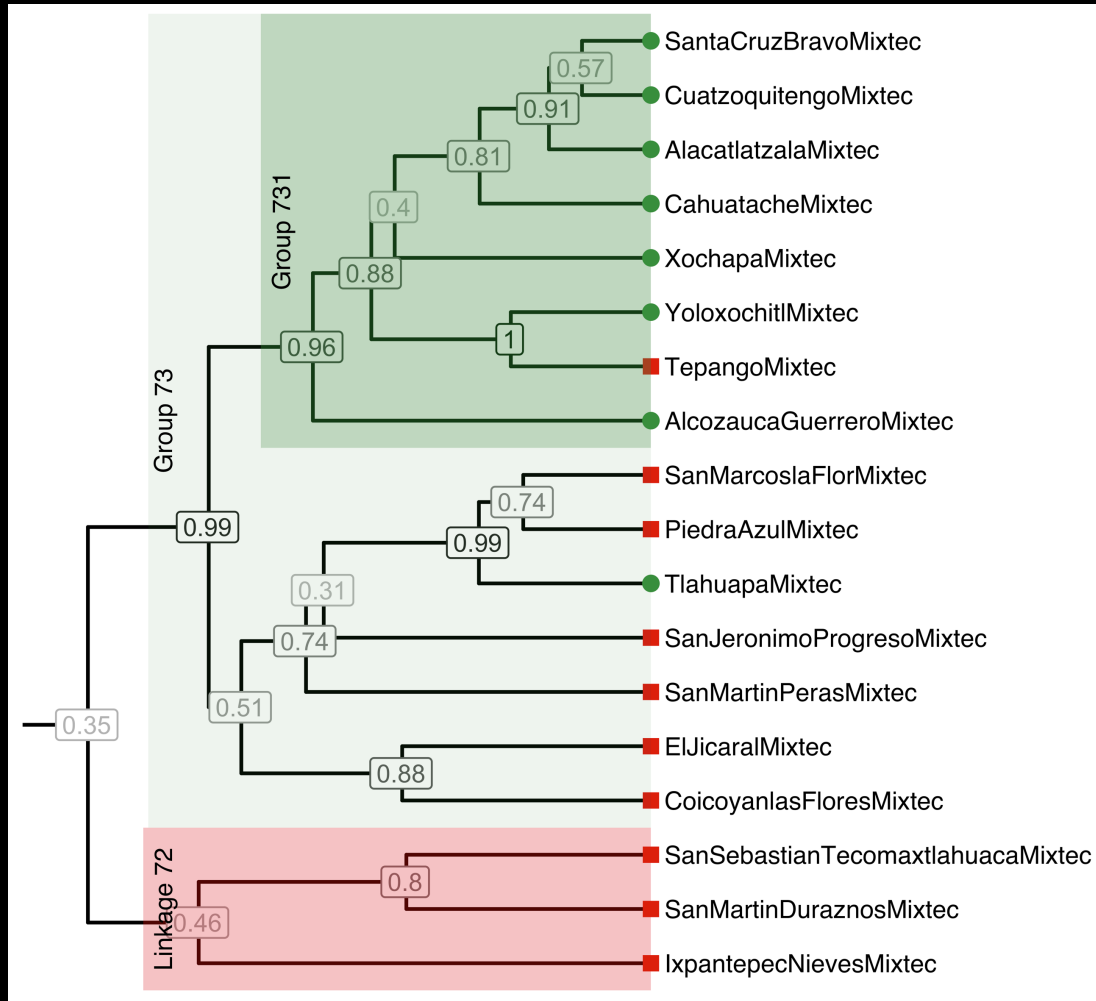
Grupo 7



7.2 y 7.3



7.2 y 7.3



Para el futuro

Mixteco:

integrar los verbos
reconstrucción de tono/cambio de
tono
encuesta lingüística actualizada
materiales para las comunidades

Integración con arqueología y antropología:

movimientos de población
redes comerciales
organización comunitaria
...

Lingüística histórica y filogenética

ir más allá del indoeuropeo y el
austronesio
más diversidad = mejores modelos
depende fundamentalmente de una
buena documentación

Colaboradores

Eric W. Campbell
UCSB

Carmen Hernández Martínez
Mixteco de San Martín Duraznos

Fidel Hernández Mendoza
Triqui de Chicahuaxtla

Christian T. DiCanio
SUNY Buffalo
Itunyoso Triqui

Rey Castillo Garcia & Jonathan Amith
Mixteco de Yoloxóchitl

Simon J. Greenhill
MPI-SHH

Griselda Reyez Basurto
Mixteco de Tlahuapa

Jeremías Salazar
Mixteco de Yucunani

Denise Kühnert
MPI-SHH tide

Yésica Ramirez
Mixteco de La Batea

Inî G. Mendoza & Simon L. Peters
Mixteco de Piedra Azul

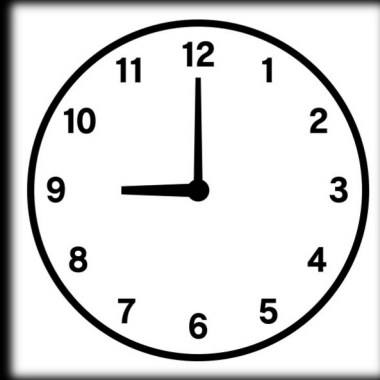
Konstantin Hoffmann
MPI-SHH tide

Juvenal Solano
Mixteco de San Sebastian del Monte

JN Martín
Mixteco de El Jicaral

Diapositivas adicionales de los
componentes del modelo

Modelos de reloj



reloj estricto

- ❖ tasa evolutiva única
- ❖ cada rama del árbol evoluciona al mismo ritmo



reloj relajado

- ❖ las tasas pueden variar de una rama a otra
- ❖ las tasas se extraen de una distribución (con parámetros estimados a partir de los datos)

Puntos de calibración

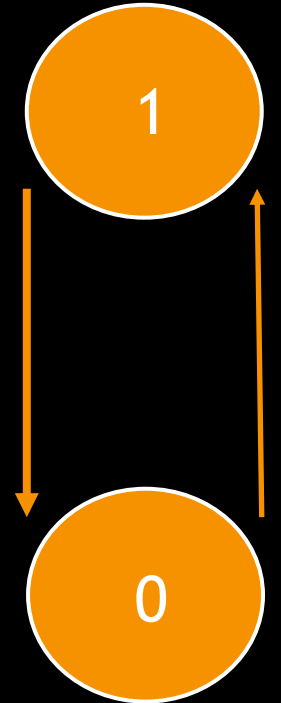
- ❖ documentos de misioneros dominicos de la época colonial
 - ❖ Antonio de los Reyes: Arte en lengua mixteca (1593)
 - ❖ Francisco de Alvarado: Vocabulario en lengua misteca (1593)
- ❖ documentos de alrededor del siglo xix
 - ❖ Francisco Belmar: Ensayo sobre la lengua trike (1897)
 - ❖ Francisco Belmar: El cuicateco (1902)



Modelo de evolución de los cognados

Binary Covarion:

- ❖ controla la velocidad de transición entre 0 y 1
- ❖ estado 'rápido' vs. 'lento'
- ❖ estima dos parámetros:
 - factor de conmutación = frecuencia de conmutación entre 'rápido' y 'lento'
 - α = diferencia de velocidad entre los estados 'rápido' y 'lento'



Modelo de árbol

Birth-death Skyline with Serial Sample (BDSKY-S)

- ❖ tasa de nacimiento (birth rate): velocidad/tasa a la que se diferencian las lenguas
- ❖ tasa de mortalidad (death rate): tasa a la que las lenguas dejan de utilizarse
- ❖ índice de muestra: cuántas lenguas de la familia se incluyeron en la muestra

