# *Sounds of Mapudungun*

## Cuantificando la diversidad geolectal en base a rasgos fonéticos

**Scott Sadowsky**
Universidad de La Frontera
Temuco, Chile
ssadowsky@gmail.com

**Paul Heggarty**
Max Planck Institute for Evolutionary Anthropology, Leipzig, Alemania
heggarty@eva.mpg.de

V CILLI/ XVI JLLM 2014
Temuco, Chile · 22-24 de octubre de 2014

# El proyecto *Sounds of the Andean Languages*

## Objetivos

Proyecto del Departamento de Lingüística del Instituto Max Planck de Antropología Evolutiva.

- **Cuantificar la divergencia fonética** que existe dentro las distintas variedades geográficas de lenguas andinas:

  **Quechua** **Aimara** **Mapudungun**

Posiblemente otras familias lingüísticas a futuro…

# El proyecto *Sounds of the Andean Languages*

## Objetivos

- Determinar los **grados de parentesco** que existen entre las variedades geográficas de cada lengua.
- Producir una **clasificación dialectológica** de cada lengua.
- En lo posible… **reconstruir parte de la historia y prehistoria** de estas lenguas.

Casos de contacto · Migraciones · Protovariedad

# El proyecto *Sounds of the Andean Languages*

## Objetivos

- **Devolverle la mano** a la comunidad.

  **Hablantes** y **aprendices** de mapudungun
  **Profesores** de mapudungun y EIB
  **Investigadores** del mapudungun
  **Público** en general

- **Sitio web**
  - Grabaciones
  - Mapas
  - Transcripciones fonéticas
  - Análisis

# Cuantificación de divergencia

## Métodos tradicionales v/s SotAL

# Métodos **tradicionales**

## Existencia de cognados

- Unidad de análisis: **Lexema**

- Pregunta:

  ¿**Existe o no un cognado** de una determinada protoforma en las variedades lingüísticas en estudio?

- Nivel de detalle: **Binario: SÍ** o **NO**
  **0 / 1**
  **0% / 100%**

# Métodos **tradicionales**

## Existencia de cognados

**Ejemplo**
¿Existen cognados del latín *castellum* en el castellano, el italiano y el francés?

| | | |
|---|---|---|
| Castellano | *castillo* | **SÍ** |
| Italiano | *castello* | **SÍ** |
| Francés | *chateau* | **SÍ** |

Grado de divergencia: **0%**

**Luego, se repite el proceso con 100 o 200 palabras.**

Pero ¡hay **grandes diferencias** entre estas palabras!

# Métodos **tradicionales**

## Existencia de cognados

**Debilidad**

- El método tradicional no puede dar cuenta de estas diferencias —a menudo profundas— porque utiliza sólo la **información más superficial**: si existe o no un cognado.

- Desecha toda la información que proporciona la fonética.

| | |
|---|---|
| Castellano | [kah.ˈt̪i.ʝo] |
| Italiano | [kas.ˈtel.lo] |
| Francés | [ʃa.ˈto] |

# Método **SotAL**

## Información fonética a nivel de rasgos

El método del proyecto *Sounds of the Andean Languages* (en adelante: **SotAL**) utiliza la **riquísima información fonética** que contiene cada palabra para cuantificar mejor la divergencia entre lenguas y variedades de lengua.

# Método **SotAL**

## Información fonética a nivel de rasgos

- Unidad de análisis: **Rasgo fonético**

  **Si algo se puede representar con AFI, se cuantifica.**

- Pregunta:

¿Cuál es la **distancia fonética entre los cognados de cada protoforma** en las variedades lingüísticas en estudio?

- Nivel de detalle: **Continuo**

  **0,00000 ⟷ 1,00000**

# Método **SotAL**

## Trabajo de campo

- **Grabar** a un gran número de hablantes provenientes de zonas geográficas diversas.

- **Elicitar** 225 palabras a cada hablante.

# Método **SotAL**

# Equipo

## Trabajo de campo

**María José Aninao**
(UFRO)

**Scott Sadowsky**
(UFRO)

**Isabel Cayunao Nahuelcheo**
(UFRO)

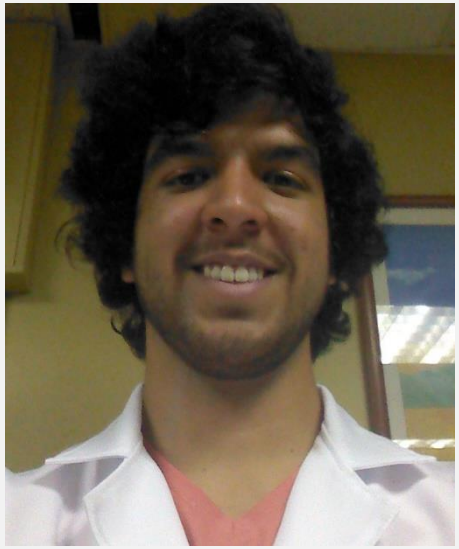
**Paul Heggarty**
(MPI-EVA)

# Método **SotAL**

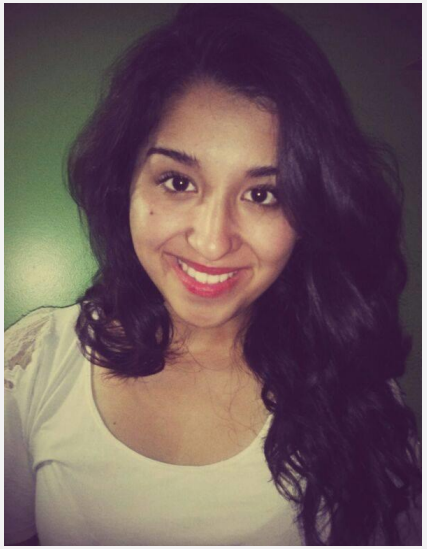
## Equipo

### Postprocesamiento y etiquetado

**Ignacia Fuentes Burgos**
(UFRO)



**Erik Farías Castro**
(UFRO)



**Margaret Mora Ortega**
(UFRO)

**Catalina Sandoval Muñoz**
(UFRO)

# Método **SotAL**

## Equipo

**Análisis fonético**

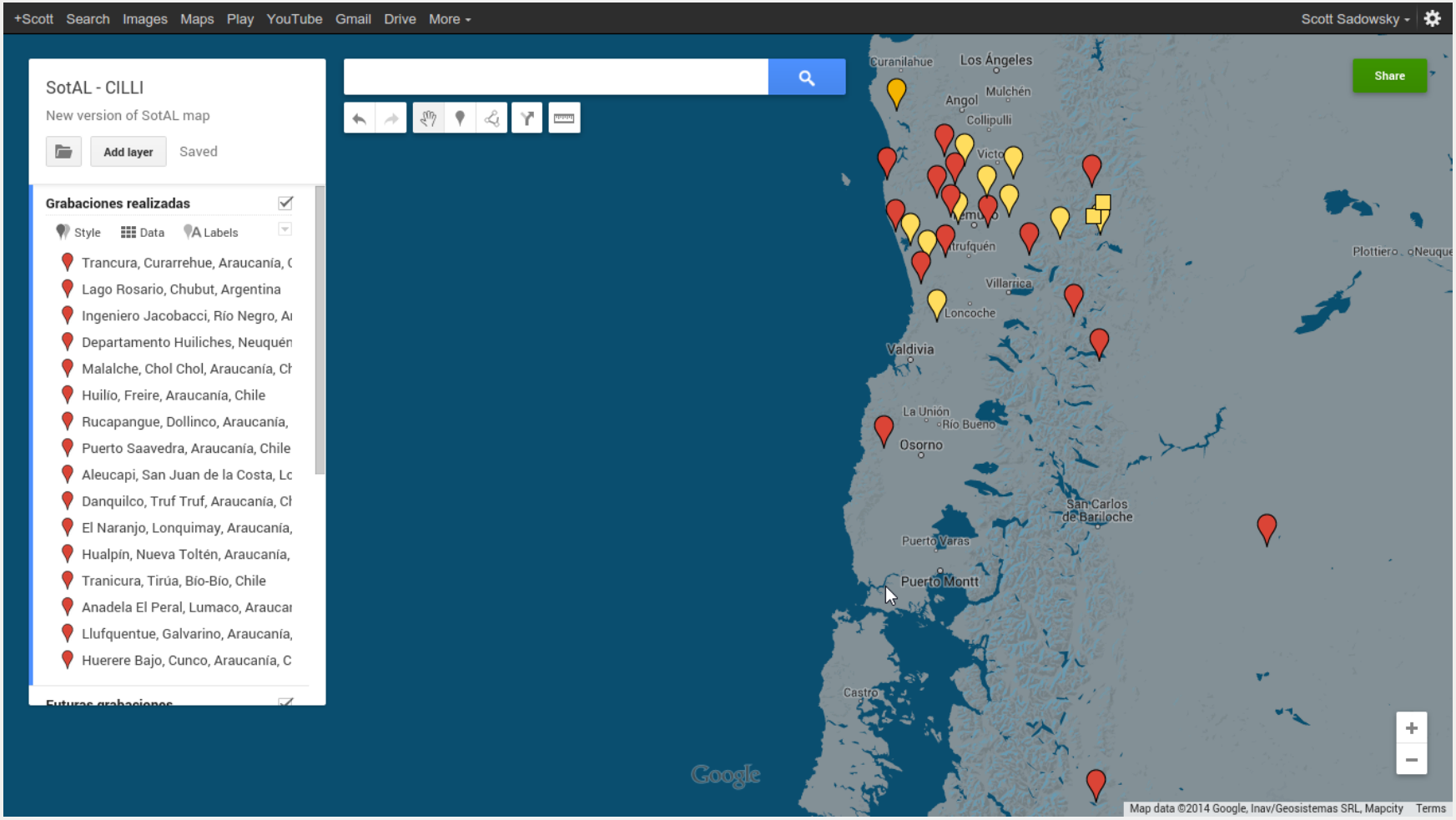
Scott Sadowsky (UFRO)

**Software / Análisis de distancias**

Paul Heggarty (MPI-EVA)

# Método **SotAL**

# Cobertura geográfica actual y futura



# Método **SotAL**

## Análisis fonético

- **Transcribir** las grabaciones fonéticamente, de manera muy estrecha.

⟨pütokon⟩

[f̬t̬˥ ˩t̬ö˩˧kön]

# Método **SotAL**

## Cuantificación de divergencia

- **Cuantificar** la divergencia fonética que existe entre los cognados de las variedades de mapudungun de las distintas localidades.

Para estos efectos, se utiliza ***Brian***, un programa computacional desarrollado por Heggarty.

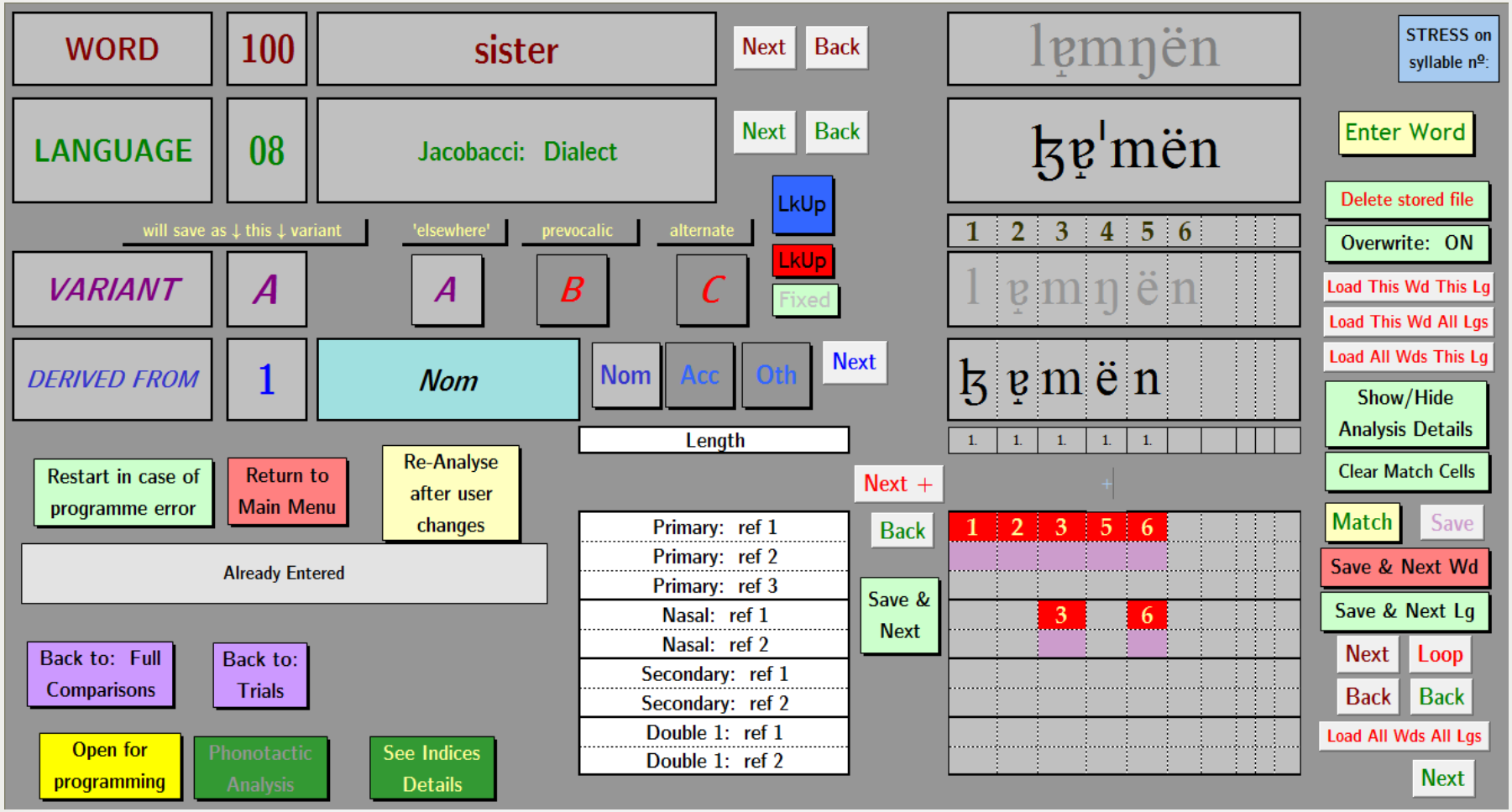
# Método **SotAL**

## Datos fonéticos

| | Trancura | Dpto. Huiliches | Lago Rosario | Ing. Jacobacci | Malalche | Huilío | Ruca-pangue | Puerto Saavedra |
|---|---|---|---|---|---|---|---|---|
| **kütral** | kʰtʲ͡ʈ͡ʂʲl | kə̟͡ʈ͡ʂʲl | kʰə˔͡ʈ͡ʂʲl | kʰə˔͡ʈ͡ʂʲl | kʰə̟˔͡ʈ͡ʂʲl | kə˔͡ʈ͡ʂʲl | kə˔͡ʈ̣͡ʂʲl | kə˔͡ʈ̣͡ʂʲl |
| **kura** | kʊˈɻᵝ | kʊˈɻᵝ | kʊˈɻᵝ | kʊˈɻᵝ | kʊˈɻᵝ | ˈkʊɻᵝ | kʊˈzᵝ | ˈkʊzᵝ |
| **mawün'** | mᵚˈwən | m̃ᵚˈwən | mᵚˈwən | mᵚˈwən | mᵚˈwən | mᵚˈwən | mᵚˈwən | mᵚˈwən̪ |
| **tromü** | ͡ʈ͡ʂᵚwˈmə | ͡ʈ͡ʂöˈmə | ͡ʈ͡ʂöˈmə | ͡ʈ͡ʂöˈməl | ͡ʈ͡ʂöˈmə | ͡ʈ͡ʂöˈmə | ʈöˈmə | ͡ʈ̣͡ʂömə |
| **kürüf** | kʰəˈɻəf | kəˈɻəf | gəˈɻʊf | kəˈɻəf | kʰəˈɻəf | kʰəˈlʊf | kʰəˈɻəf | kəˈɻəf |
| **kofke** | köfˈkë | köf̟ˈcë | köfˈcʰë | köfˈcʰël | ˈkʰöfcʰë | köfˈcë | köfˈcë | ˈköfcʰë |
| **chadi** | ͡tʃᵝˈθɪ | ͡tʃᵝˈθɪ | ͡tʃᵝˈθɪ | ͡tʃᵝˈθɪ | ͡tʃᵝˈθɪ | ͡tʃᵝˈθɪ | ʃᵝˈθɪ | ˈʃᵝθɪ |
| **iwiñ** | jɪˈwɪɲ | jɪˈwɪɲ | ɪˈwɪɲ | ɪˈwɪɲ | ʝɪˈwɪɲ | ɪˈwɪɲ | ʒɪˈwɪɲ | ɪˈwɪɲ |

# Método **SotAL**

# Determinación de correlatos entre fonos



# Método **SotAL**

## ¿Cuánta divergencia entre cada variedad?

| | | 01 | 02 | 03 | 04 | 05 | 06 | 07 | 08 | 09 |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Proto-Mapudungun | Chol Chol: Dialect | Dollinco: Dialect | Saavedra: Dialect | Freire: Dialect | Curarrehue: Dialect | Dpto Huiliches: Dialect | Jacobacci: Dialect | Lago Rosario: Dialect |
| === | ========== | ==== | ==== | ==== | ==== | ==== | ==== | ==== | ==== | ==== |
| | 'SISTER' | | | | | | | | | |
| 01 | Proto-Mapudungun | PMpg | 0.729 | 0.828 | 1.000 | 1.000 | 1.000 | 1.000 | 0.732 | 0.724 |
| 02 | Chol Chol: Dialect | | Chol | 0.798 | 0.727 | 0.727 | 0.727 | 0.727 | 0.911 | 0.461 |
| 03 | Dollinco: Dialect | | | Doll | 0.828 | 0.828 | 0.828 | 0.828 | 0.807 | 0.553 |
| 04 | Saavedra: Dialect | | | | PtSv | 1.000 | 1.000 | 1.000 | 0.732 | 0.724 |
| 05 | Freire: Dialect | | | | | Frei | 1.000 | 1.000 | 0.732 | 0.724 |
| 06 | Curarrehue: Dialect | | | | | | Cura | 1.000 | 0.732 | 0.724 |
| 07 | Dpto Huiliches: Dialect | | | | | | | Huil | 0.732 | 0.724 |
| 08 | Jacobacci: Dialect | | | | | | | | Jaco | 0.468 |
| 09 | Lago Rosario: Dialect | | | | | | | | | LRos |

# Método **SotAL**

## Cálculo de divergencia fonética

Los valores de distancia fonética entre **cada fono** de **cada palabra** de **cada variedad de la lengua** son los datos que se analizan en la siguiente etapa.

Veamos un **ejemplo**: la realización de ⟨**lamngen**⟩ en hablantes de mapudungun de ocho localidades:

Trancura (Curarrehue)
Lago Rosario (Chubut)
Malalche (Chol Chol)
Rucapangue (Dollinco)
Dpto. Huiliches (Neuquén)
Jacobacci (Río Negro)
Huilío (Freire)
Puerto Saavedra

# Métodos **tradicionales**

## ¿Cognados? Sí. Divergencia = 0

⟨lamngen⟩ → **SÍ**
⟨lamngen⟩ → **SÍ**
⟨lamngen⟩ → **SÍ**
⟨lamngen⟩ → **SÍ**
⟨lamngen⟩ → **SÍ**
⟨lamngen⟩ → **SÍ**
⟨lamngen⟩ → **SÍ**
⟨lamngen⟩ → **SÍ**

# Método **SotAL**

## Palabras fonéticamente distintas = 5

[ lɐ̟mˈŋën ]

[ lɐ̟mˈŋën ]

[ jɐ̟mˈŋë ]

[ ɮɐ̟ˈmën ]

[ lɐ̟mɪˈën ]

[ lɐ̟mˈŋën ]

[ lɐ̟mˈwën ]

[ lɐ̟mˈŋën ]

# Método **SotAL**

## Número de fonos divergentes = 6

| | | | | | |
|---|---|---|---|---|---|
| l | ɐ̣ | m | ŋ | ë | n |
| l | ɐ̣ | m | ŋ | ë | n |
| j | ɐ̣ | m | ŋ | ë | ∅ |
| ɮ | ɐ̣ | m | ∅ | ë | n |
| l | ɐ̣ | m | ɪ | ë | n |
| l | ɐ̣ | m | ŋ | ë | n |
| l | ɐ̣ | m | w | ë | n |
| l | ɐ̣ | m | ŋ | ë | n |

# Método **SotAL**

## ¿Cuánta divergencia entre cada fono?

| | | | | | |
|---|---|---|---|---|---|
| 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 |
| 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 |
| 0,67 | 1,00 | 1,00 | 1,00 | 1,00 | 0,00 |
| 0,79 | 1,00 | 1,00 | 0,00 | 1,00 | 1,00 |
| 1,00 | 1,00 | 1,00 | 0,16 | 1,00 | 1,00 |
| 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 |
| 1,00 | 1,00 | 1,00 | 0,90 | 1,00 | 1,00 |
| 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 |

# Método **SotAL**

## Cuantificación de divergencia (lista)

| | | 01 | 02 | 03 | 04 | 05 | 06 | 07 | 08 | 09 |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Proto-Mapudungun | Chol Chol: Dialect | Dollinco: Dialect | Saavedra: Dialect | Freire: Dialect | Curarrehue: Dialect | Dpto Huiliches: Dialect | Jacobacci: Dialect | Lago Rosario: Dialect |
| 01 | Proto-Mapudungun | PMpg | 0.835 | 0.844 | 0.861 | 0.852 | 0.835 | 0.828 | 0.819 | 0.836 |
| 02 | Chol Chol: Dialect | | Chol | 0.890 | 0.900 | 0.899 | 0.892 | 0.886 | 0.872 | 0.886 |
| 03 | Dollinco: Dialect | | | Doll | 0.905 | 0.890 | 0.893 | 0.871 | 0.874 | 0.881 |
| 04 | Saavedra: Dialect | | | | PtSv | 0.914 | 0.900 | 0.893 | 0.890 | 0.903 |
| 05 | Freire: Dialect | | | | | Frei | 0.905 | 0.901 | 0.883 | 0.903 |
| 06 | Curarrehue: Dialect | | | | | | Cura | 0.893 | 0.884 | 0.908 |
| 07 | Dpto Huiliches: Dialect | | | | | | | Huil | 0.864 | 0.896 |
| 08 | Jacobacci: Dialect | | | | | | | | Jaco | 0.870 |
| 09 | Lago Rosario: Dialect | | | | | | | | | LRos |

# Análisis de divergencia

## Métodos tradicionales v/s SotAL

# Métodos **tradicionales**

## Análisis: Divisiones binarias

- Modelo **arbóreo**
- Única relación posible: **división binaria (bifurcación)**
- Ejemplo: **Croese (1980)**

# Métodos **tradicionales** en dialectología

## Análisis: Divisiones binarias

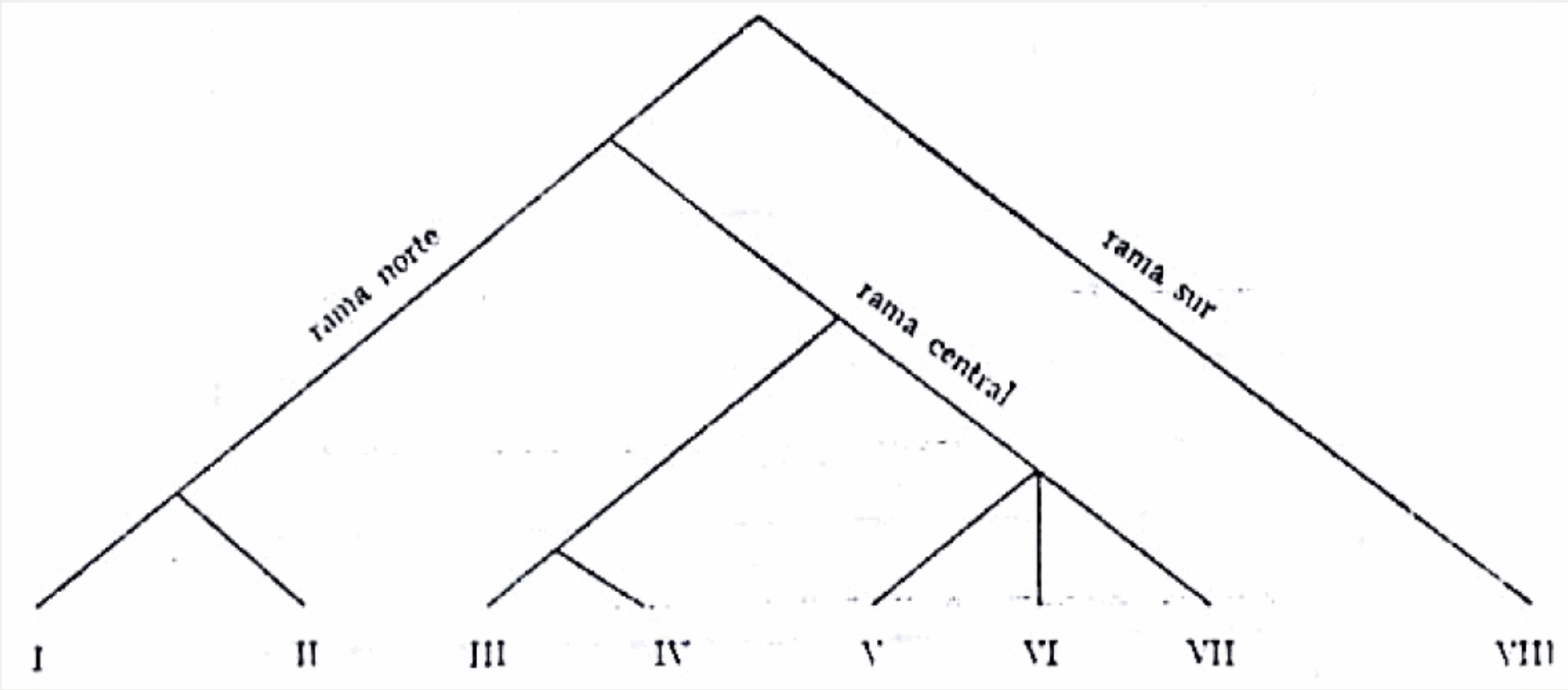



Diagrama arbóreo de variedades de mapudungun según Croese (1980).

# Métodos **tradicionales** en dialectología

## Debilidades del modelo arbóreo

- **Todos es bifurcación.** Supone que las poblaciones humanas siempre se dividen **binariamente**: en dos grupos… luego dos grupos más… luego dos más.

- **No** puede dar cuenta de los **continuos dialectales**.
- **No** puede dar cuenta del **contacto lingüístico**.
- **No** puede dar cuenta de **cambios paralelos**.

# Método SotAL

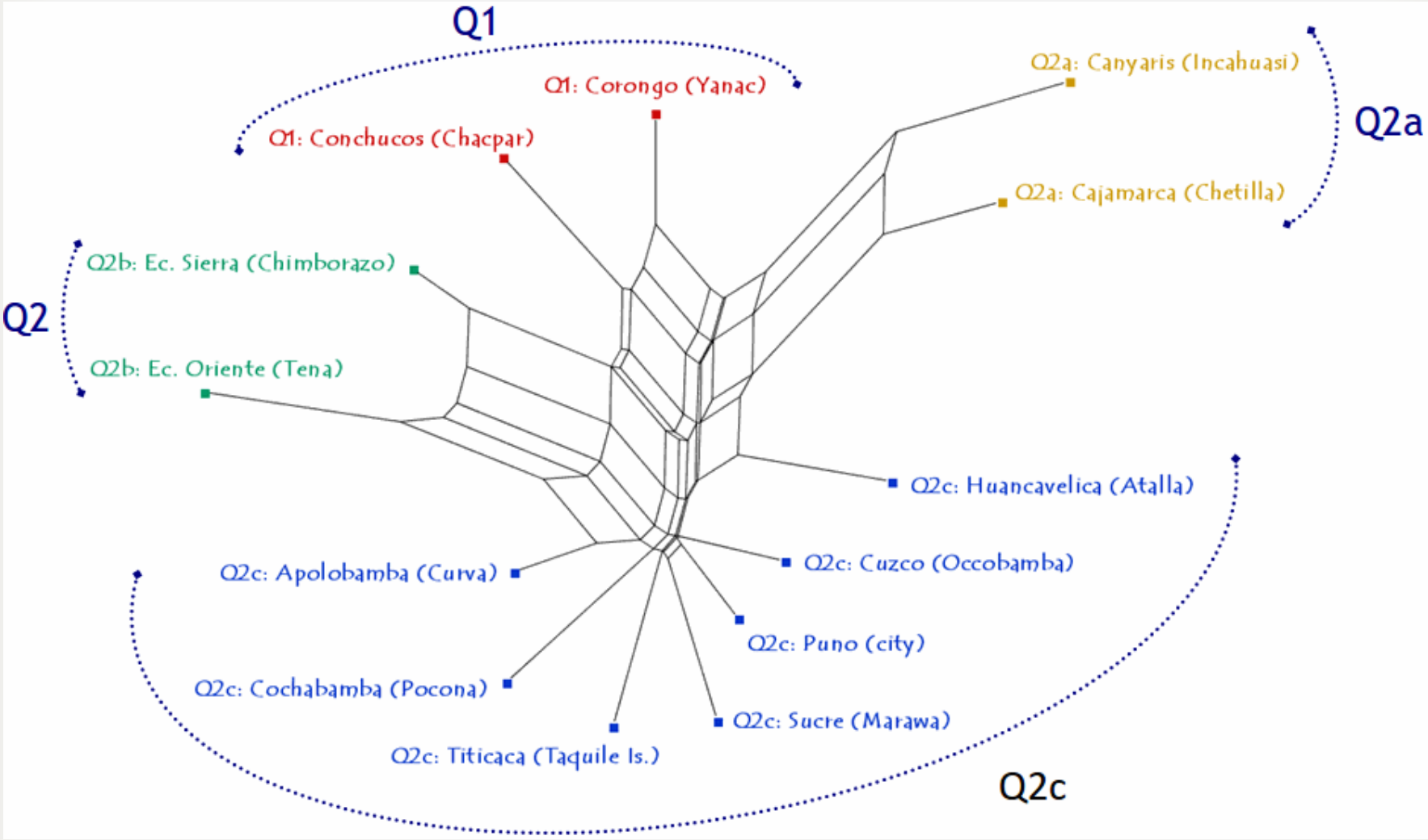
## Modelo de redes filogenéticas

- Modelo desarrollado en la **bioinformática** para representar información **genética** en base al cálculo de distancias (grados de divergencia).

- Permite modelar mucho mejor la **evolución compleja de las lenguas** y variedades de lengua.

# Método **SotAL**

## Quechua: Red filogenética



# Método **SotAL**

## Análisis: Redes complejas

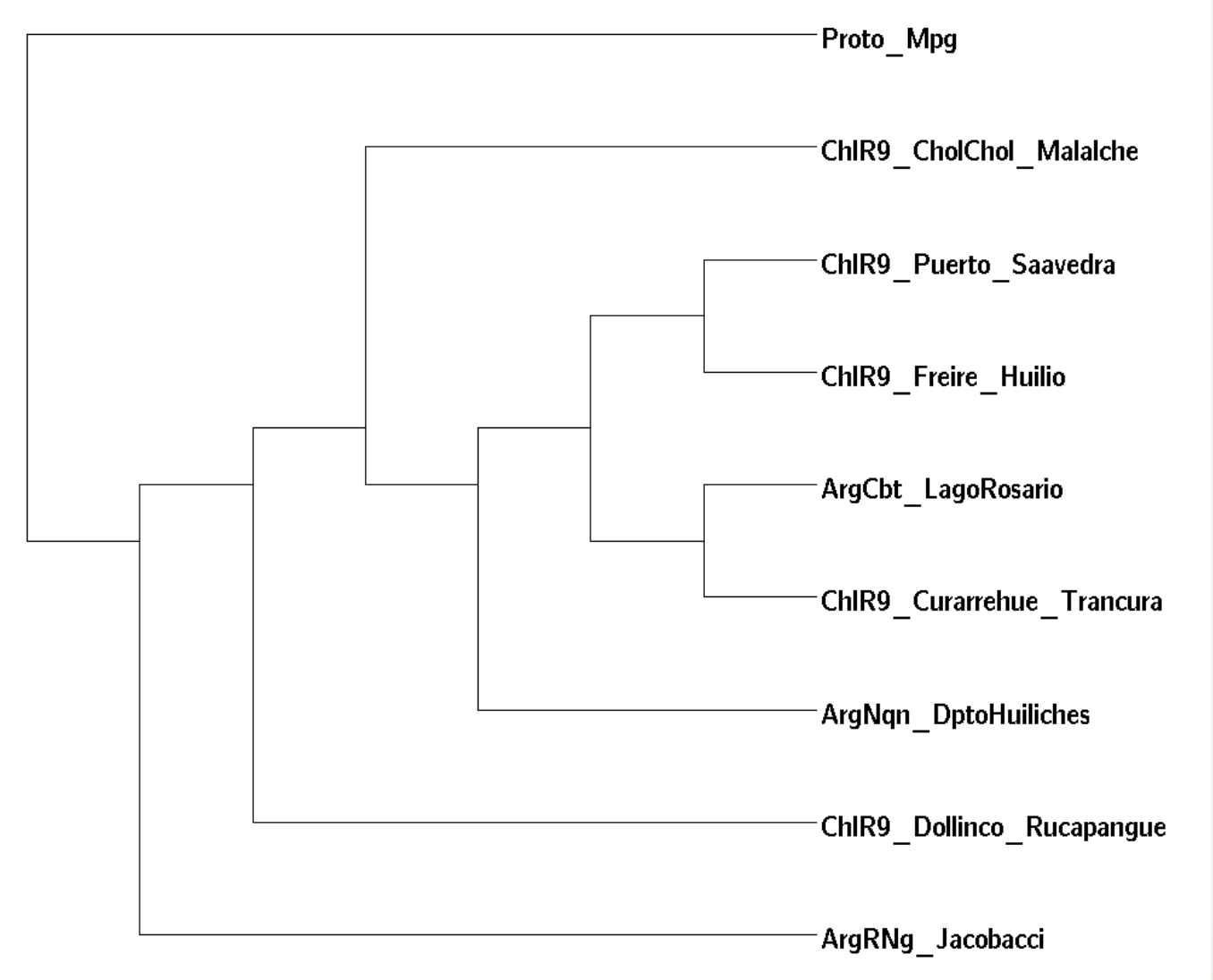
**¡ADVERTENCIA!**

Los **resultados** que vamos a ver a continuación ni siquiera **ni siquiera pueden considerarse provisorios**.

- Tamaño de muestra: **Sólo 8 hablantes**.
- Representatividad de muestra: **Baja**.
- Software de cuantificación fonética: Aún **no completamente adaptado** para mapudungun.
- Palabras analizadas: **Sólo 60 de 225**.

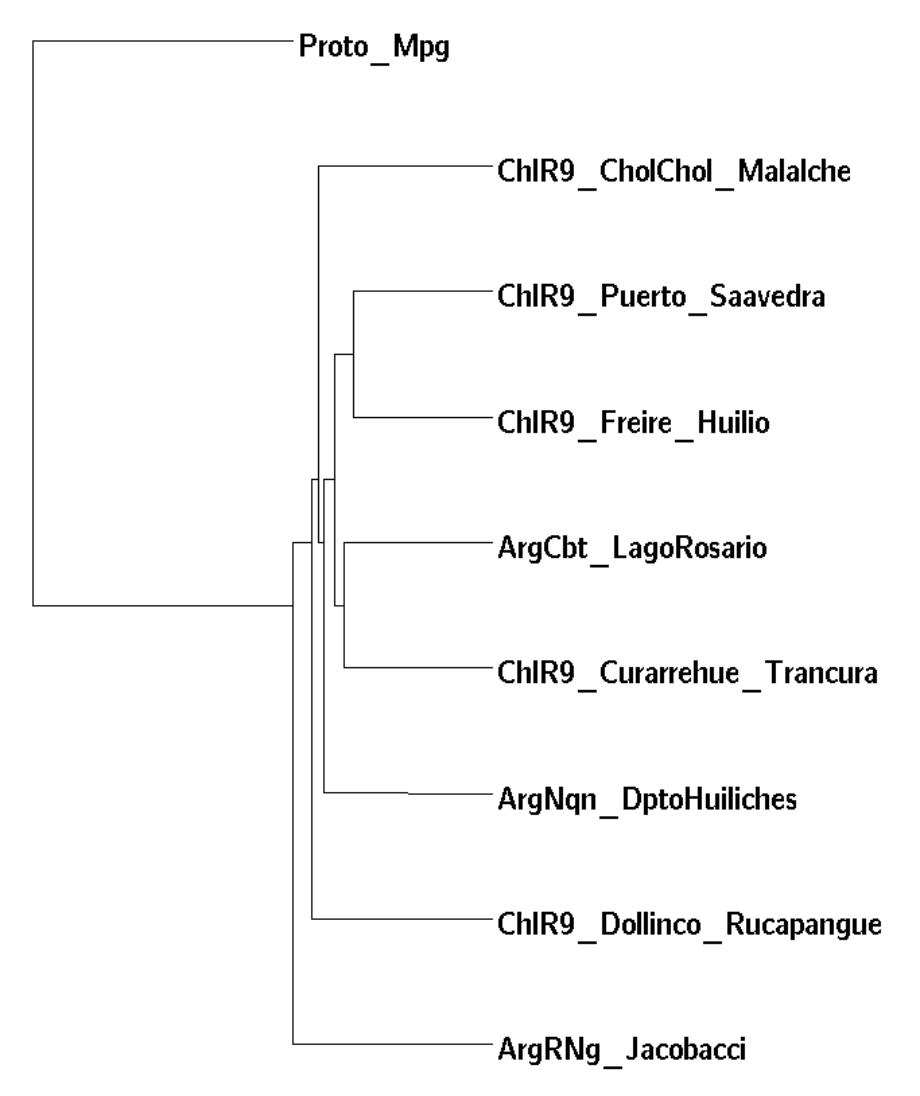
# Método **SotAL**

## Cladograma



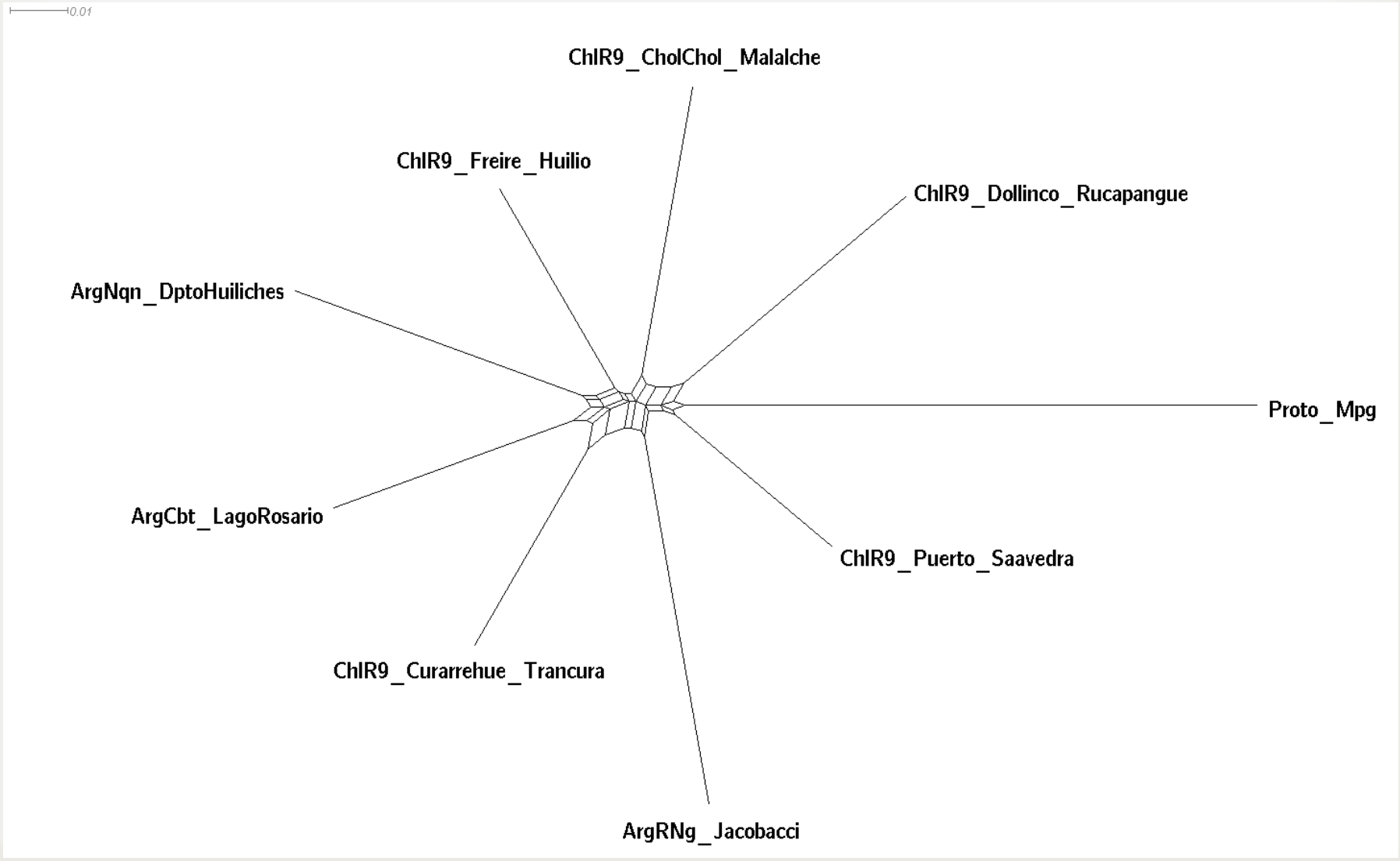
# Método **SotAL**

## **Phylograma**



# Método **SotAL**

## Red de vecinos (*Neighbor net*)



# Proyecto Sounds of the Andean Languages

## Sitio web

- El sitio web está en **marcha blanca.**
  - Afinar **detalles** (presentación visual, funcionalidad)
  - Corregir **errores** (transcripción, buscador, estructura de datos)
  - Arreglar **errores** de programación (base de datos, interfaz)
  - Y sobre todo… **¡incorporar a más hablantes**!
- Por eso, **todavía no está disponible en internet.**
- Pero vamos a ver un **anticipo** ahora…

# *Sounds of Mapudungun*

## Cuantificando la diversidad geolectal en base a rasgos fonéticos

**Scott Sadowsky**
Universidad de La Frontera
Temuco, Chile
ssadowsky@gmail.com

**Paul Heggarty**
Max Planck Institute for Evolutionary Anthropology, Leipzig, Alemania
heggarty@eva.mpg.de

CILLI/JLLM 2014
Temuco, Chile · 22-24 de octubre de 2014