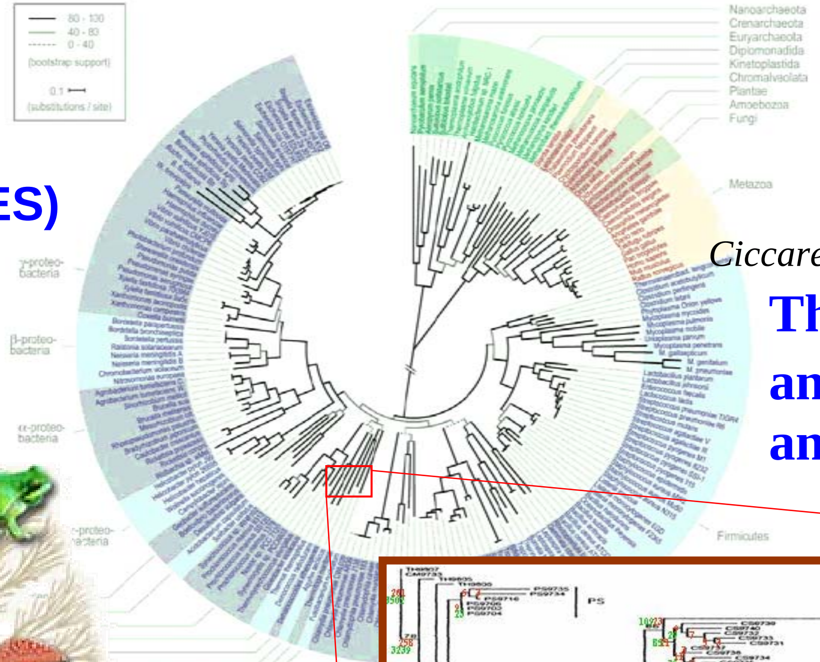
# Genetic Similarity Networks
## in populations and in metapopulations

Emilio Hernández-García[1], Alejandro F. Rozenfeld[2], Sophie Arnaud-Haond[3;4], Víctor M. Eguíluz[1], Ester Serrão[3], Carlos M. Duarte[2]

1. **IFISC**, Instituto de Física Interdisciplinar y Sistemas Complejos (CSIC-UIB),Palma de Mallorca, Spain
2. **IMEDEA**, Instituto Mediterráneo de Estudios Avanzados (CSIC-UIB), Esporles, Mallorca, Spain
3. **CCMAR**, CIMAR-Laboraório Associado, Universidade do Algarve, Faro, Portugal
4. **EEP/LEP**-Laboratoire Environnement Profond, IFREMER Centre de Brest, France

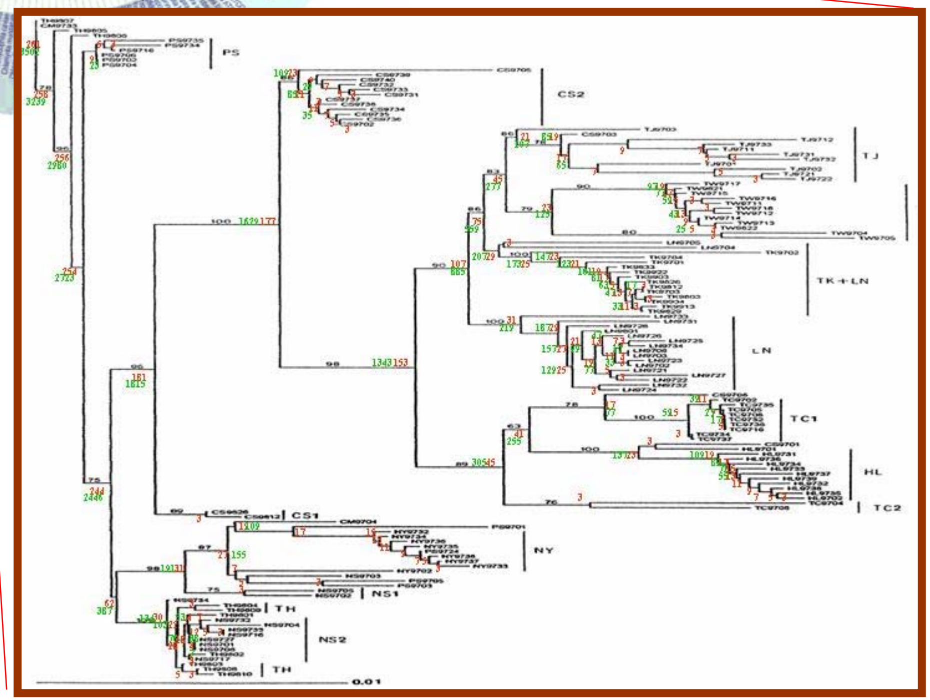
# NETWORKS (TREES) in EVOLUTION

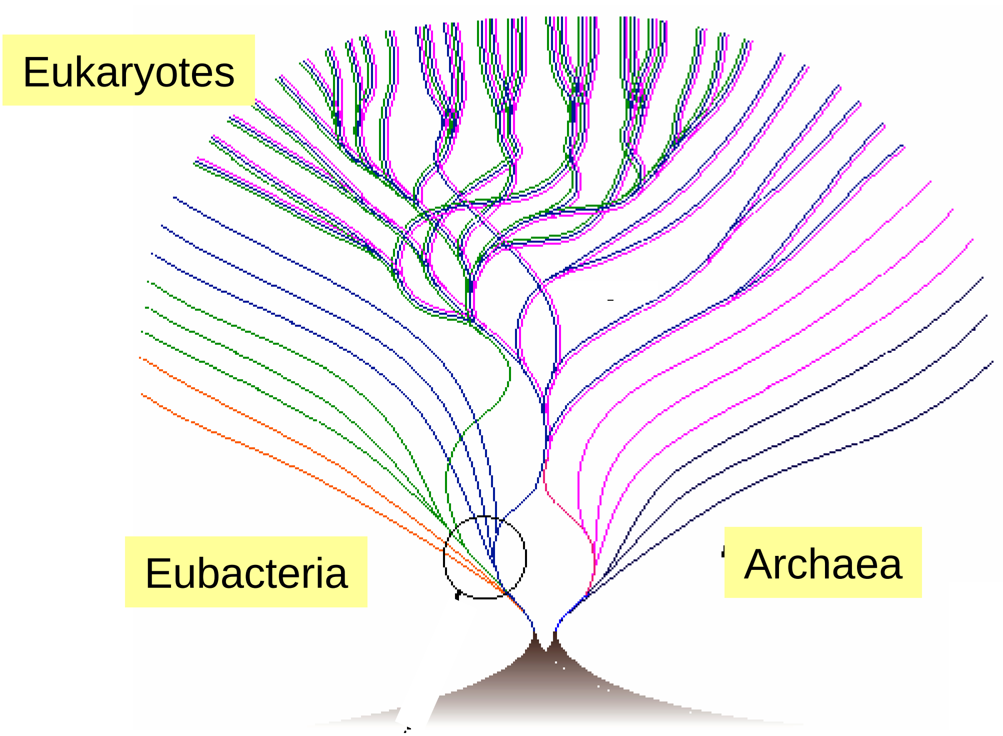



*Ciccarelli, F. D. et al. Science (2006)*

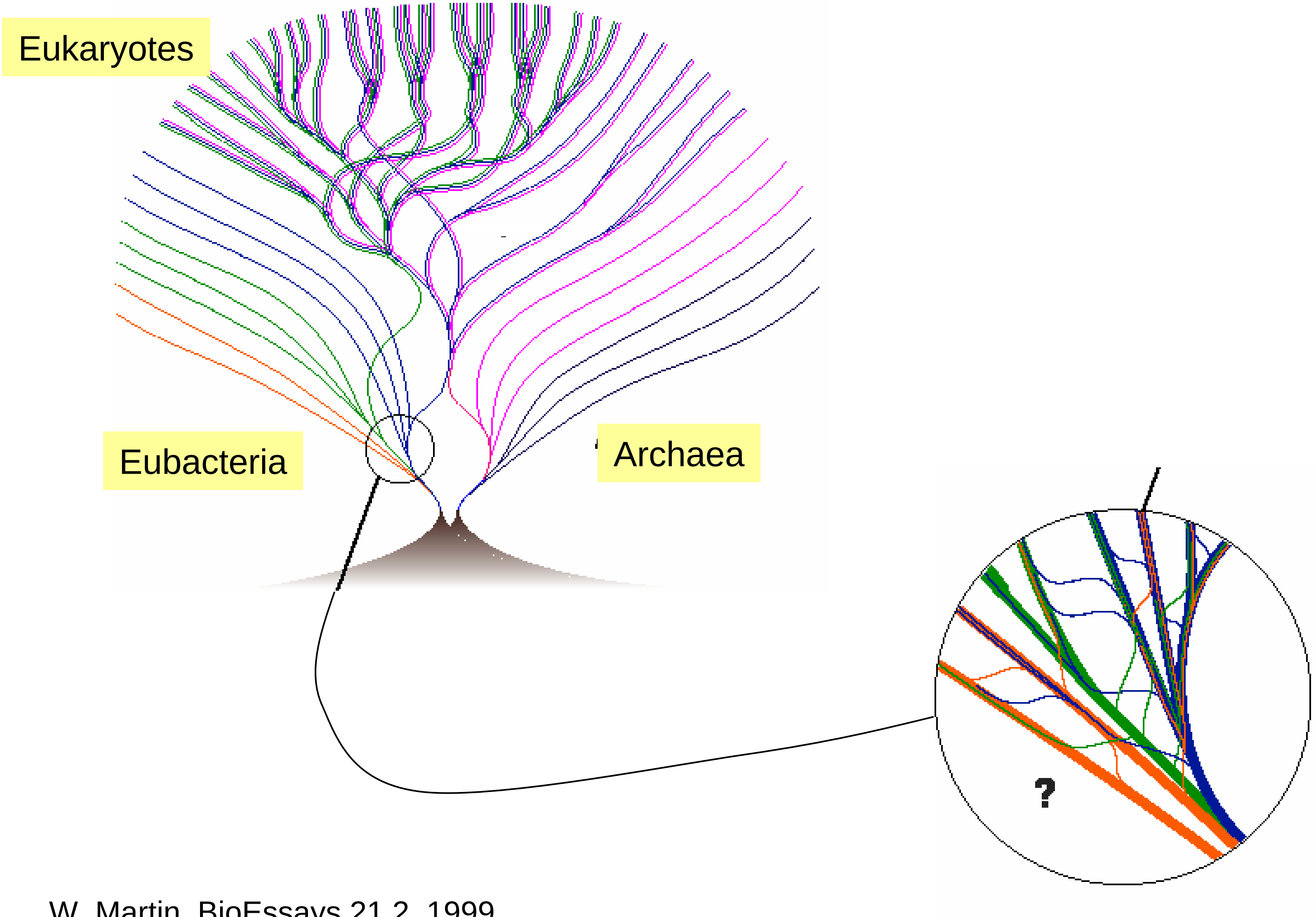
## The "Tree of Life" and phylogenetic trees and networks

Tree of Life from **tolweb.org**

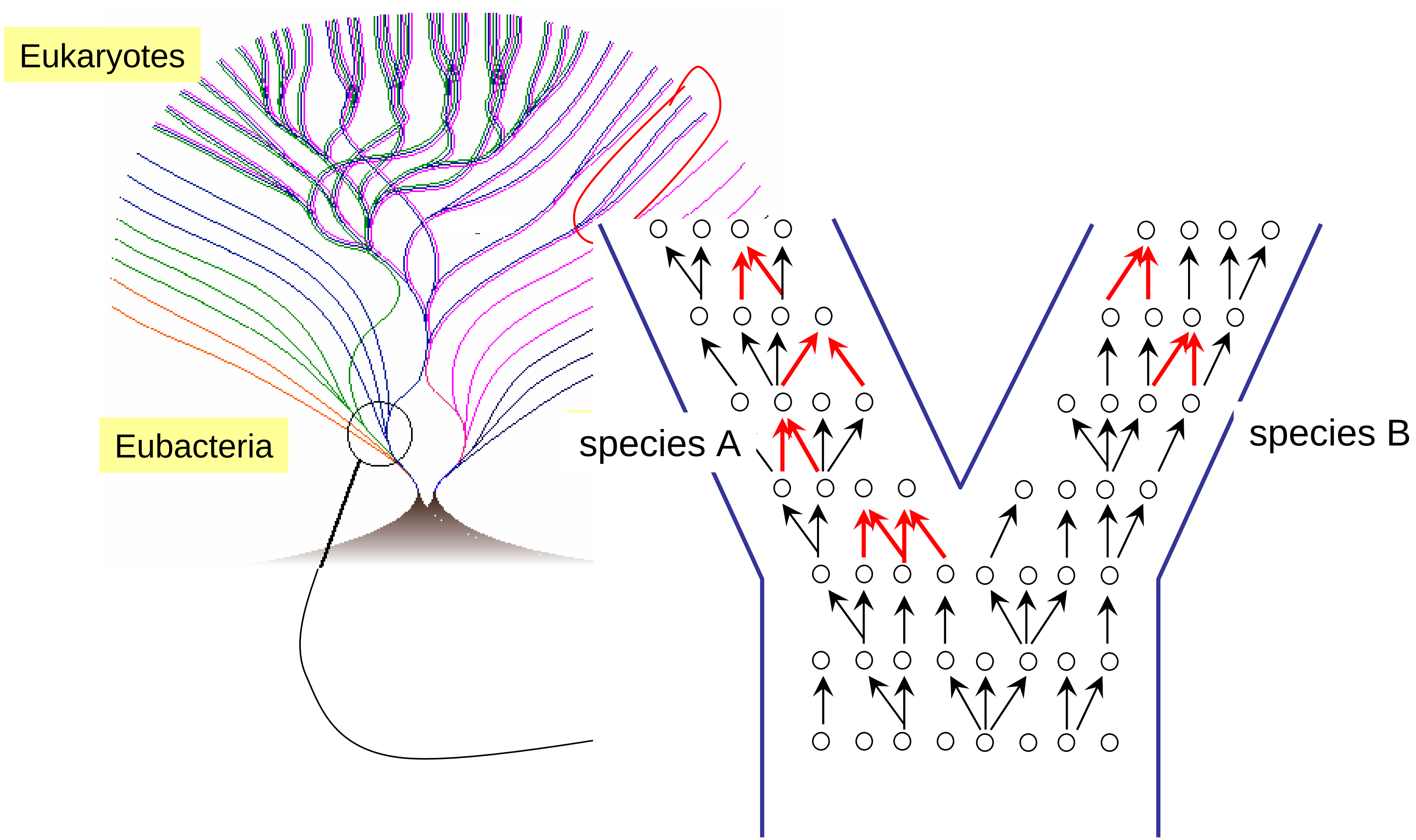


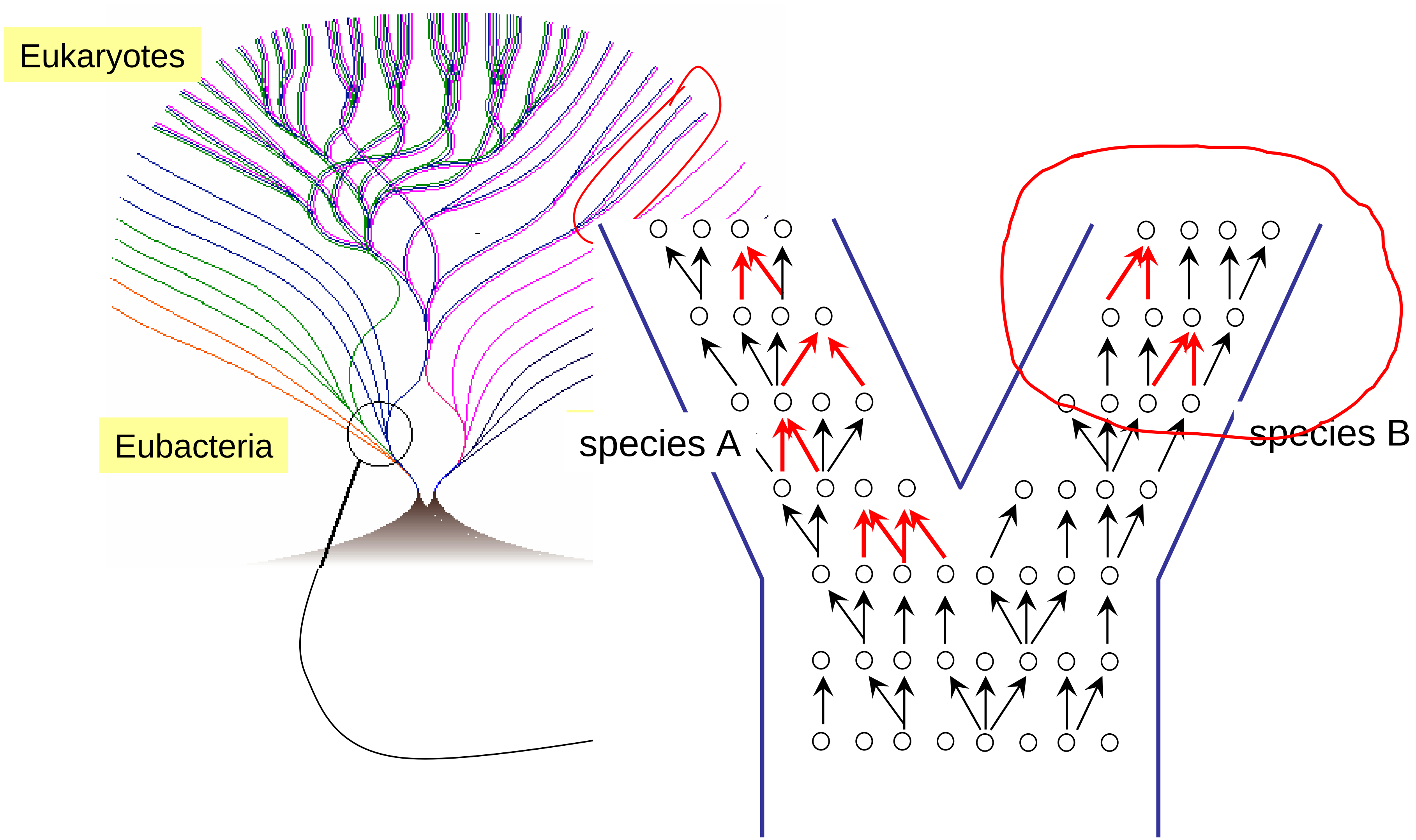



W. Martin, BioEssays 21.2, 1999




W. Martin, BioEssays 21.2, 1999

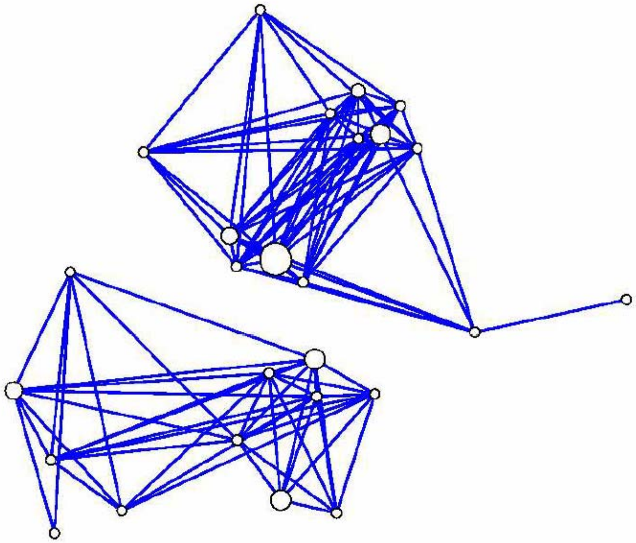

Eukaryotes
Eubacteria
species A
species B

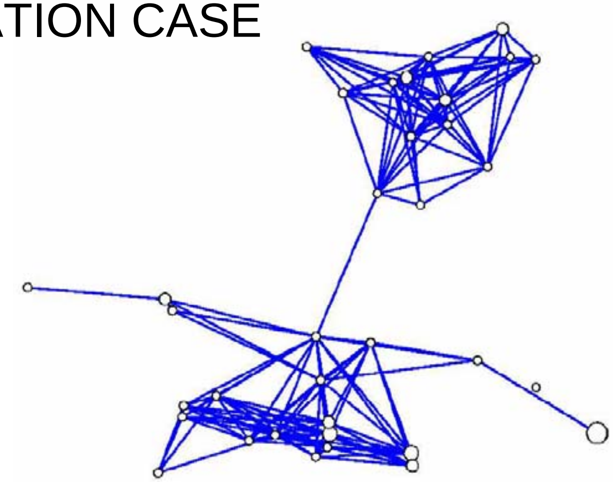

Eukaryotes
Eubacteria
species A
species B

## NETWORKS OF GENETIC SIMILARITY

INTRAPULATION CASE

Rodalquilar population





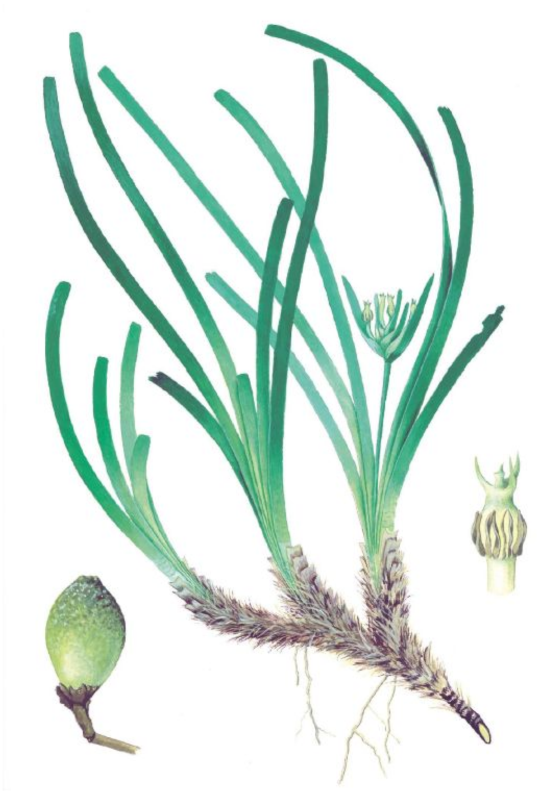
Roquetas population

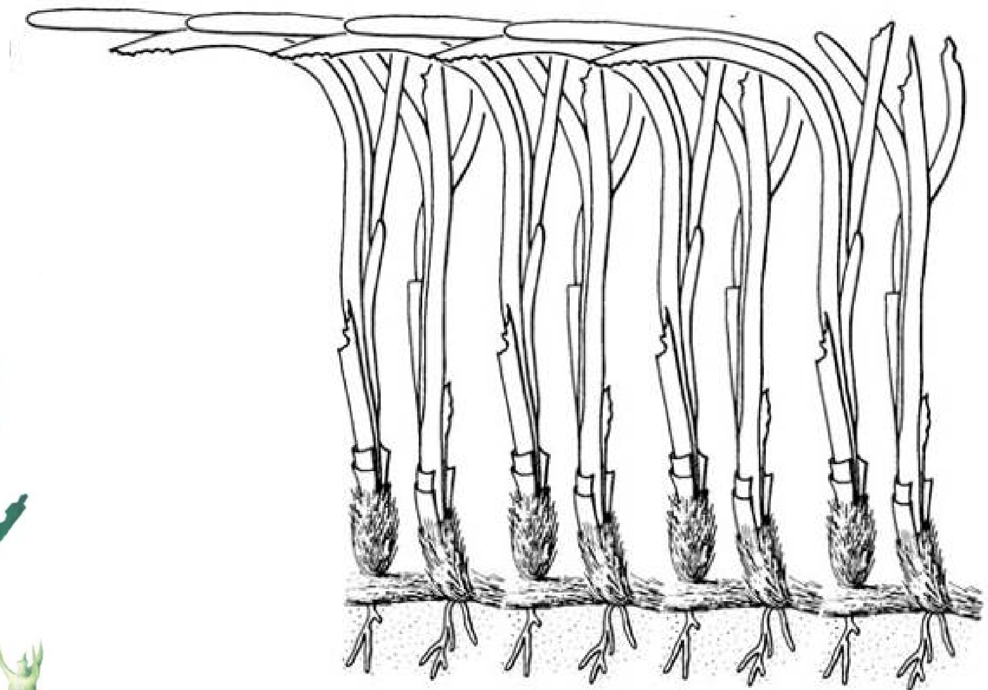
THE AIM:

- To provide biologists with novel methods for analysis of population structure based on modern network theory advances
- To enrich the science of networks with a new case study with distinct properties

*Posidonia oceanica* meadows in Cabrera (Balearic islands)

Posidonia oceanica meadows

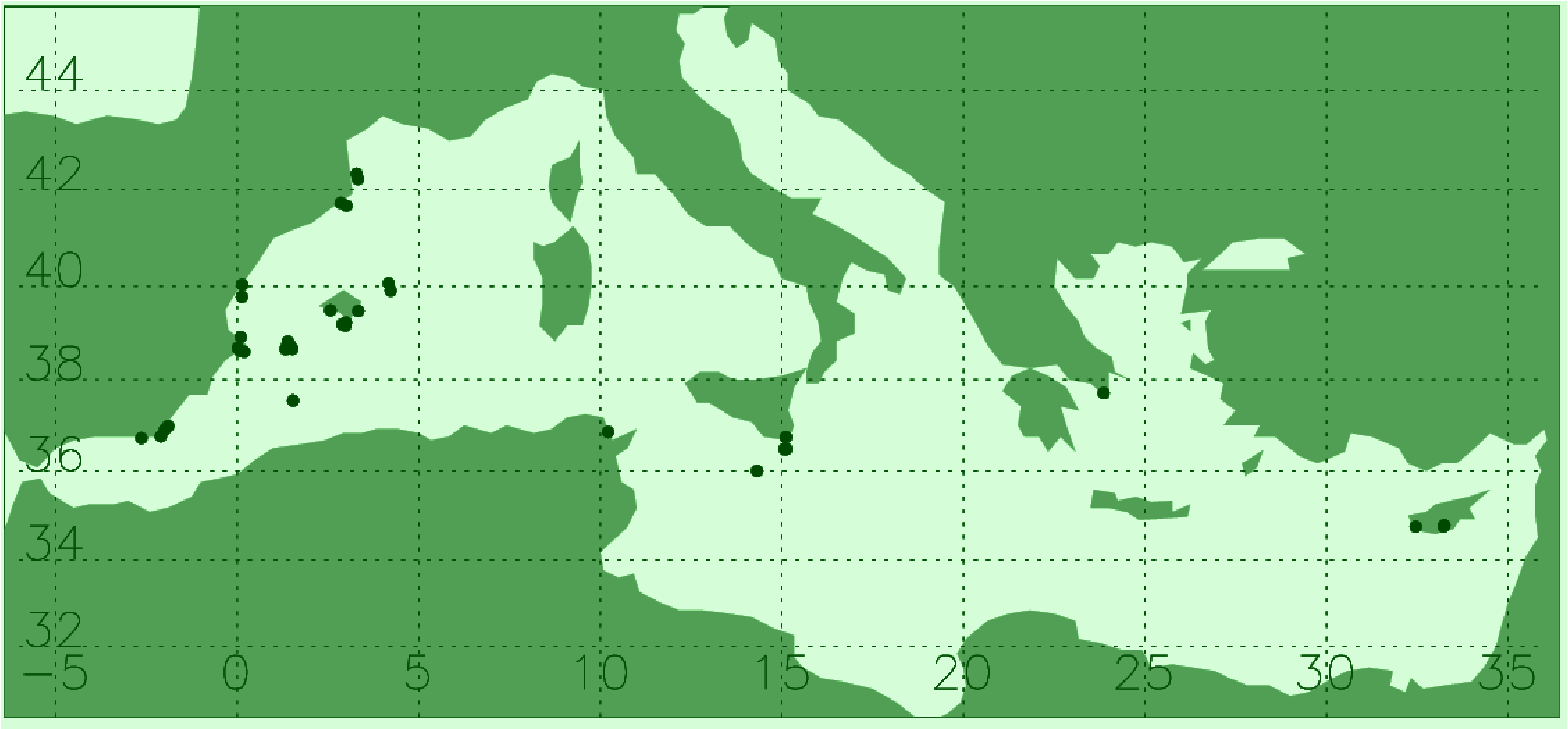




Combined clonal + sexual modes of reproduction

# Constructing networks of genetic similarity in *Posidonia oceanica*

- Sampling the individuals: 40 ramets collected from 40 populations located in different places of the Mediterranean
- Genotyping: Characterization by 7 microsatellite loci.
- Network construction:
  - Distance definition: Distance matrix
  - Network construction by thresholding the distance matrix




The 40 sampling locations in the Mediterranean

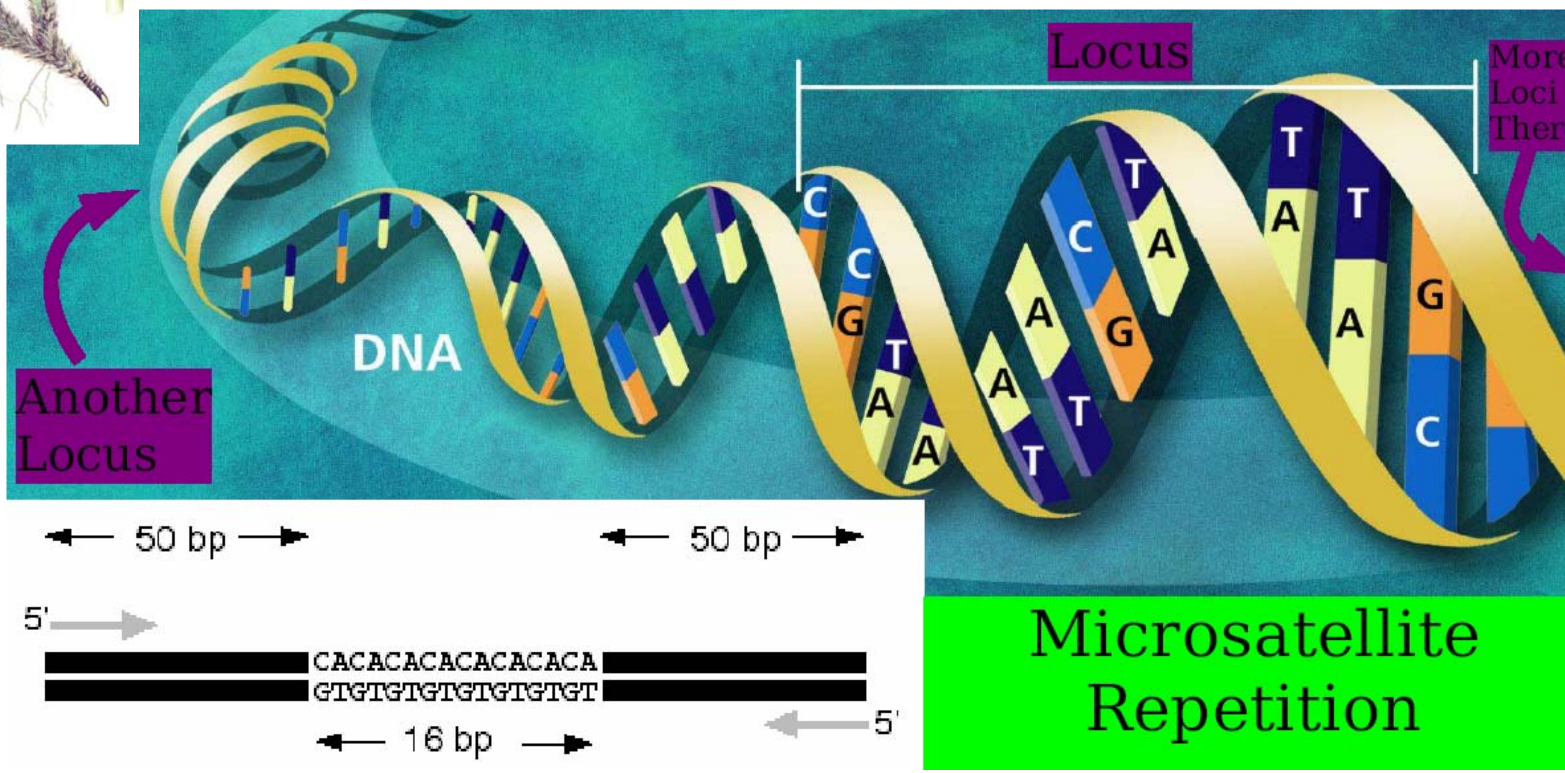
# Constructing networks of genetic similarity in *Posidonia oceanica*

- Sampling the individuals: 40 ramets collected from 40 populations located in different places of the Mediterranean
- **Genotyping**: Characterization by 7 microsatellite loci.
- Network construction:
  - Distance definition: Distance matrix
  - Network construction by thresholding the distance matrix

Genetically caracterizing the individuals:

MICROSATELLITES: hypervariable markers useful for intrapopulation studies

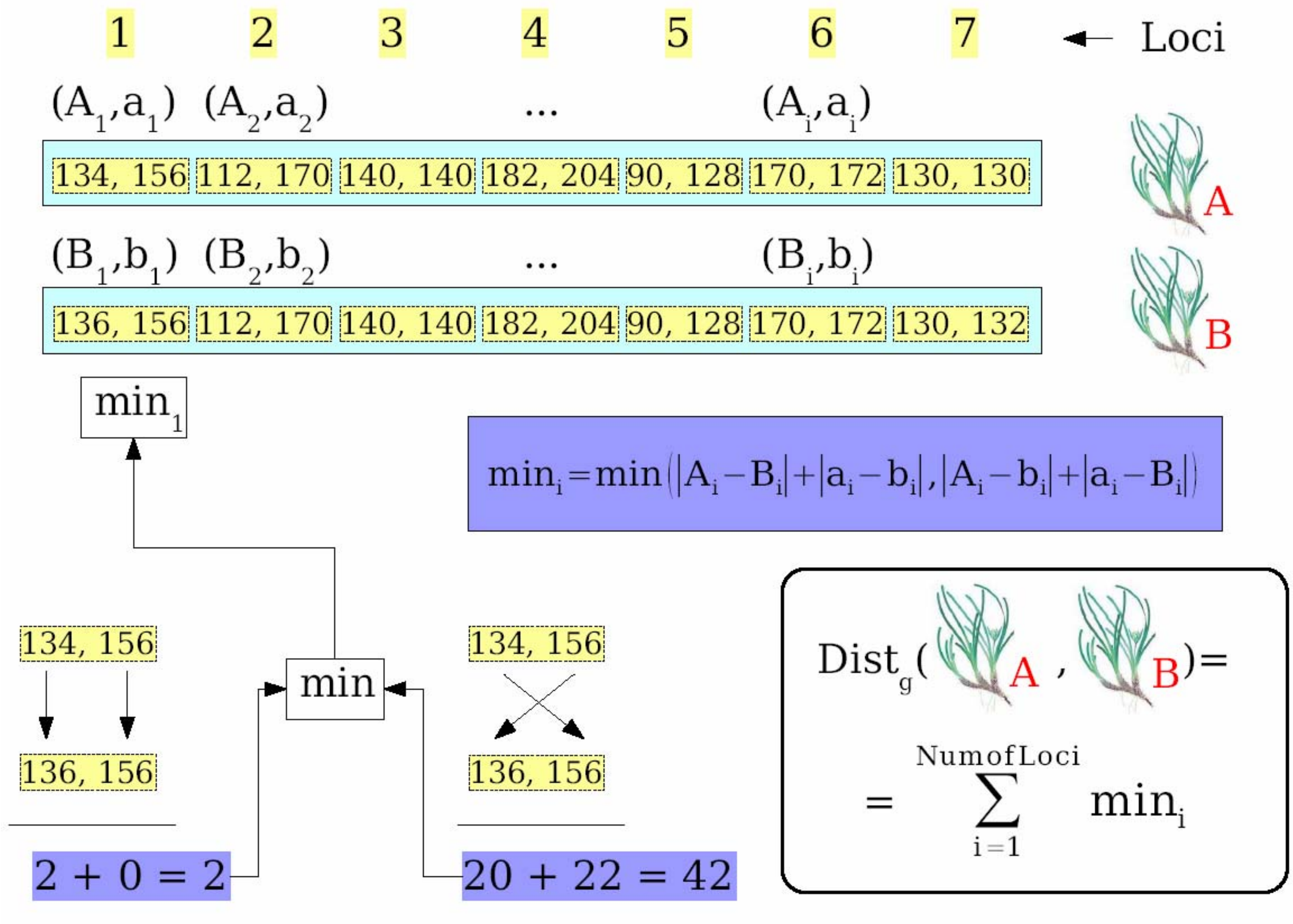
7 loci analyzed



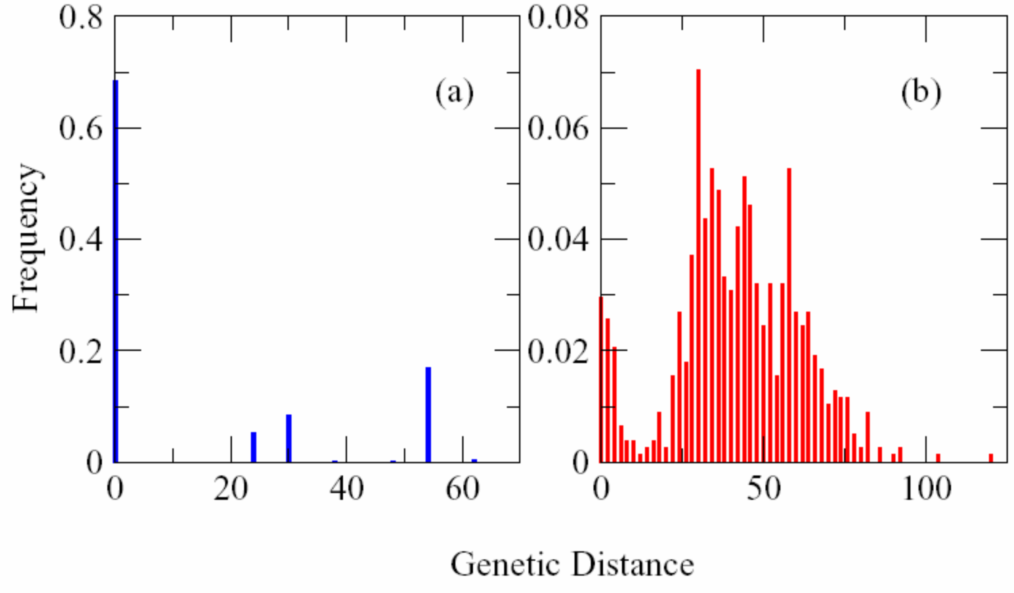
# Constructing networks of genetic similarity in *Posidonia oceanica*

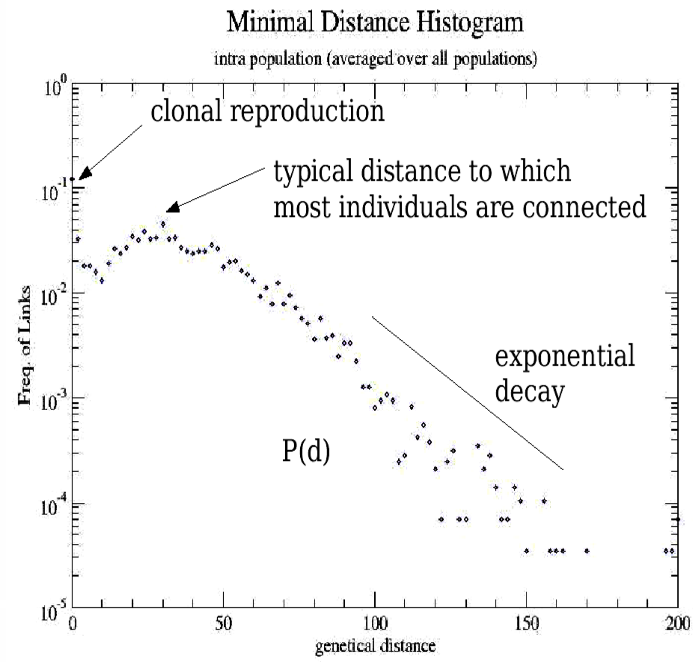
- Sampling the individuals: 40 ramets collected from 40 populations located in different places of the Mediterranean
- Genotyping: Characterization by 7 microsatellite loci.
- Network construction:
  - **Distance** definition: Distance matrix
  - Network construction by thresholding the distance matrix

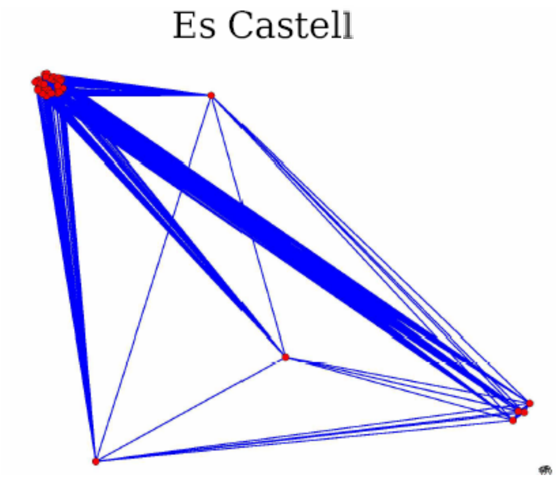
# DEFINING A DISTANCE BETWEEN INDIVIDUAL RAMETS OR GENETS

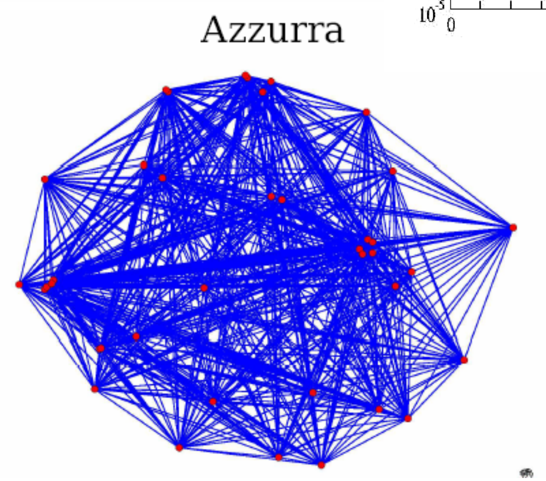


## Intrapopulation Genetic Diversity Spectra (distance histograms)



# Constructing networks of genetic similarity in *Posidonia oceanica*

- Sampling the individuals: 40 ramets collected from 40 populations located in different places of the Mediterranean
- Genotyping: Characterization by 7 microsatellite loci.
- Network construction:
  - Distance definition: Distance matrix
  - Network construction by **thresholding** the distance matrix

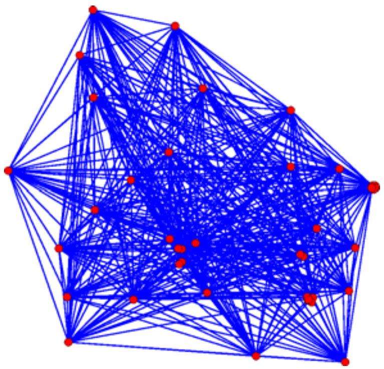
## NETWORK GENERATION:

Links between ramets genetically closer than a threshold distance $D_{th}$

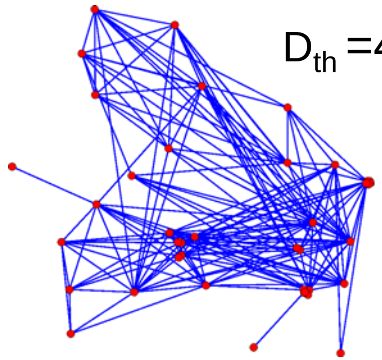
Examples: Interactions in a stock market (Onnela et al, 2003)

Functional networks in the human brain (Eguíluz et al, 2005).

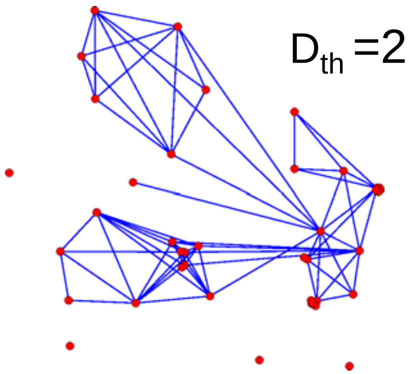
full



$D_{th}$ =40

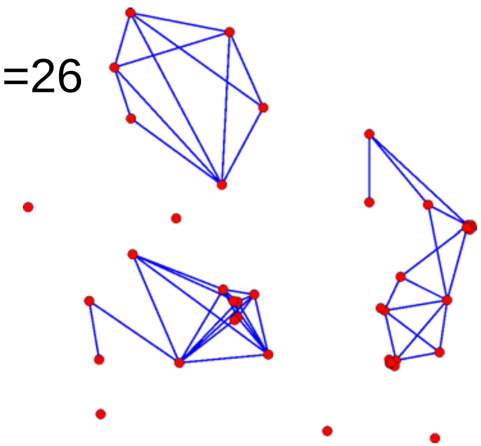


$D_{th}$ =29

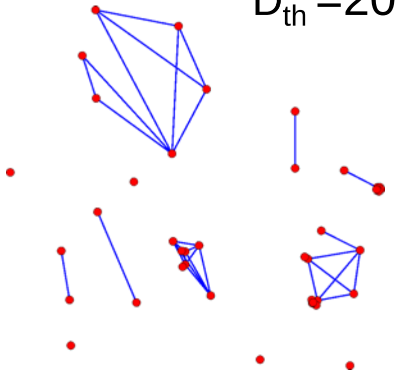


**Azzura5 population**

$D_{th}$ =26



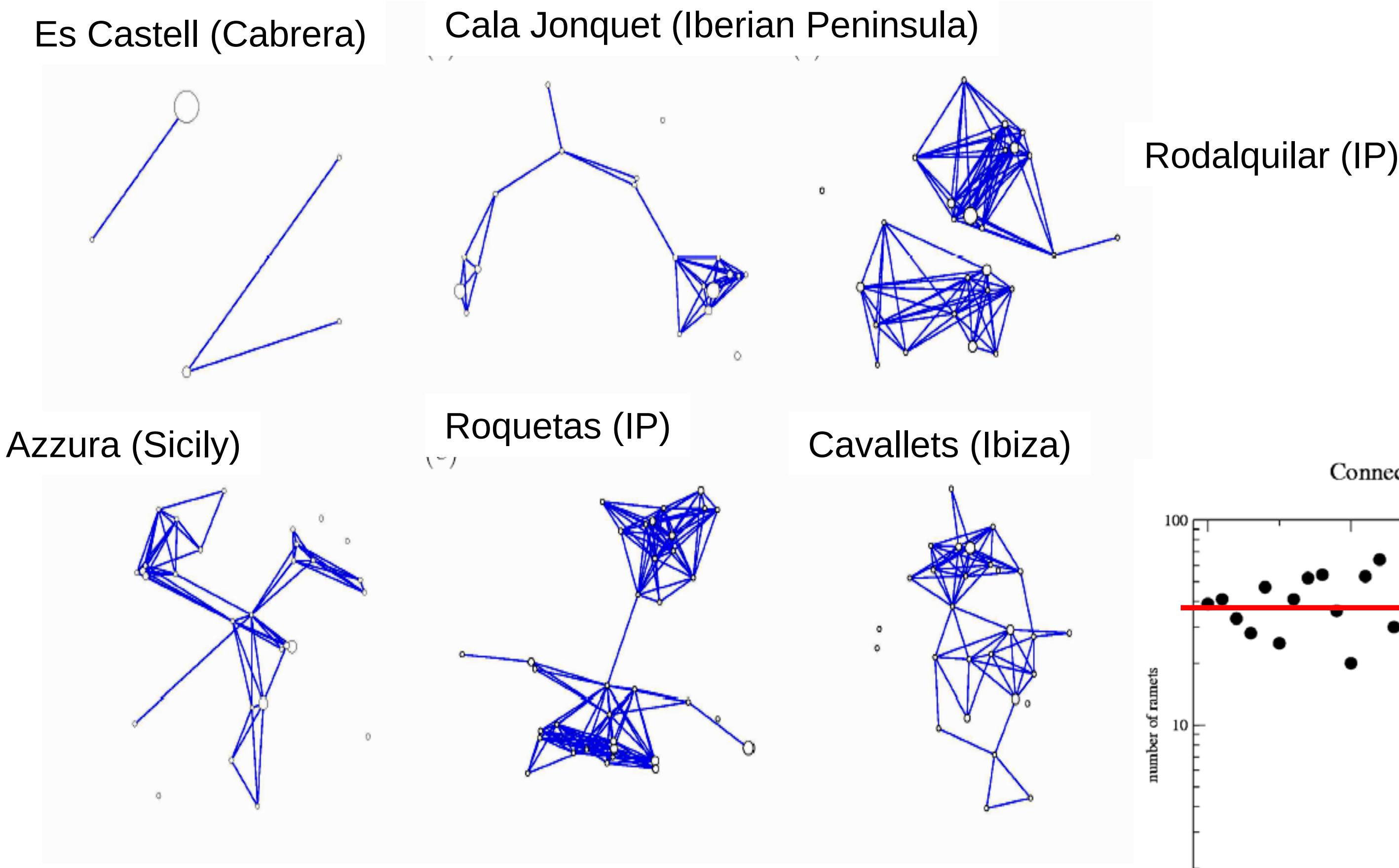
$D_{th}$ =20

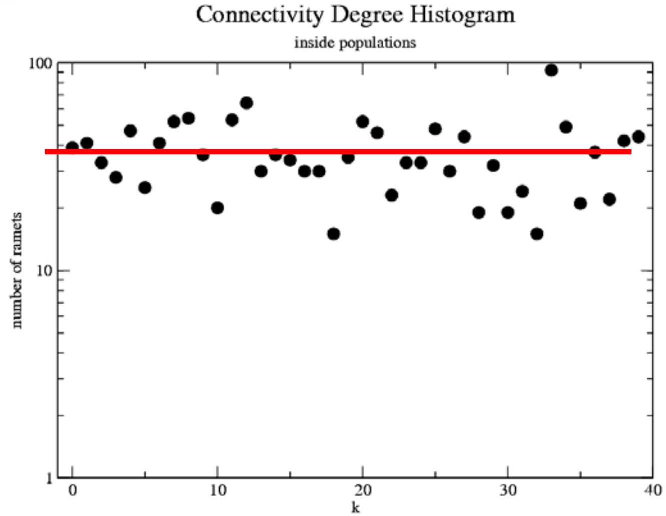


In all populations fragmentation occurs when $D_{th} \approx$ expected distance parents-offspring $\delta$. This is close to the nonclonal maximum in the distance distribution.

Thus studying the networks for $D_{th} \approx \delta$ is a rather natural choice.

# GENET NETWORKS FOR $D_{th} \approx \delta$ (expected distance parent-offspring)




Groups, clustering, …→ population structure

IFISC

| | G | S | C | L | Cr | Lr | ⟨k⟩ | ρ |
|---|---|---|---|---|---|---|---|---|
| Aqua Azzura 3 | 31 | 28 | 0.80 | 3.13 | 0.22 | 2.17 | 4.71 | 0.16 |
| Aqua Azzura 5 | 29 | 25 | 0.78 | 2.43 | 0.26 | 2.01 | 4.76 | 0.17 |
| Addaia | 25 | 17 | 0.72 | 1.64 | 0.41 | 1.91 | 5.60 | 0.23 |
| Amathous 3 | 18 | 11 | 0.77 | 1.45 | 0.33 | 1.86 | 4.67 | 0.27 |
| Amathous 5 | 25 | 24 | 0.73 | 3.14 | 0.29 | 2.19 | 4.56 | 0.19 |
| Agios Nicolaos | 28 | 18 | 0.81 | 2.41 | 0.27 | 2.13 | 5.50 | 0.20 |
| Calabardina | 40 | 40 | 0.62 | 3.44 | 0.19 | 2.25 | 5.90 | 0.15 |
| Cala Giverola | 17 | 10 | 0.84 | 1.22 | 0.70 | 1.76 | 4.35 | 0.27 |
| Cala Jonquet | 20 | 18 | 0.79 | 2.95 | 0.37 | 1.89 | 4.60 | 0.24 |
| Campomanes | 22 | 12 | 0.66 | 1.77 | 0.25 | 2.18 | 4.00 | 0.19 |
| Carboneras | 16 | 16 | 0.78 | 1.75 | 0.75 | 1.49 | 8.50 | 0.57 |
| El Arenal | 32 | 27 | 0.76 | 2.74 | 0.27 | 2.11 | 5.25 | 0.17 |
| Es Castell | 05 | 03 | 0.00 | 1.33 | 0.00 | 1.33 | 1.20 | 0.30 |
| Es Pujols | 27 | 24 | 0.74 | 2.76 | 0.57 | 1.97 | 5.70 | 0.22 |
| Es Cal des Oli | 15 | 07 | 0.73 | 1.14 | 0.49 | 1.98 | 2.80 | 0.20 |
| Ses Illetes | 21 | 20 | 0.78 | 2.62 | 0.36 | 1.82 | 5.52 | 0.28 |
| Cala Fornells | 05 | 03 | 1.00 | 1.00 | 1.00 | 1.00 | 1.20 | 0.30 |
| La Fossa Calpe | 31 | 29 | 0.63 | 2.76 | 0.24 | 2.08 | 5.55 | 0.18 |
| Las Rotes | 34 | 21 | 0.75 | 1.98 | 0.20 | 2.25 | 4.65 | 0.14 |
| Los Genoveces | 14 | 13 | 0.86 | 1.45 | 0.86 | 1.34 | 7.71 | 0.59 |
| Magaluf | 26 | 18 | 0.64 | 1.88 | 0.39 | 2.11 | 4.46 | 0.18 |
| Malta | 29 | 24 | 0.67 | 2.19 | 0.33 | 1.89 | 5.24 | 0.19 |
| Marzamemi | 31 | 28 | 0.63 | 2.48 | 0.39 | 2.07 | 5.61 | 0.19 |
| Es Port | 05 | 05 | 0.90 | 1.10 | 0.90 | 1.10 | 3.60 | 0.90 |
| Paphos | 26 | 17 | 0.74 | 1.65 | 0.43 | 1.95 | 6.54 | 0.26 |
| Playa Cavallets | 28 | 24 | 0.63 | 2.45 | 0.30 | 1.97 | 4.86 | 0.18 |
| Port Lligat | 12 | 07 | 0.95 | 1.10 | 0.43 | 1.59 | 4.17 | 0.38 |
| Porto Colom | 21 | 16 | 0.83 | 1.42 | 0.75 | 1.55 | 7.62 | 0.38 |
| Punta Fanals | 26 | 26 | 0.70 | 2.06 | 0.48 | 1.83 | 7.54 | 0.30 |
| Rodalquilar | 27 | 14 | 0.84 | 1.32 | 0.37 | 1.75 | 8.22 | 0.32 |
| Roquetas | 35 | 34 | 0.79 | 2.47 | 0.33 | 1.86 | 8.74 | 0.26 |
| C.Sta.Maria 13 | 20 | 19 | 0.73 | 2.37 | 0.43 | 1.79 | 5.50 | 0.29 |
| C.Sta.Maria 7 | 22 | 16 | 0.66 | 1.74 | 0.61 | 1.82 | 4.91 | 0.23 |
| Cala Torreta | 21 | 10 | 0.75 | 1.84 | 0.13 | 2.37 | 3.33 | 0.17 |
| Torre de la Sal | 15 | 13 | 0.71 | 1.64 | 0.58 | 1.64 | 4.93 | 0.35 |
| Tunis | 34 | 30 | 0.77 | 2.80 | 0.28 | 2.15 | 5.41 | 0.16 |
| Xilxes | 12 | 05 | 0.47 | 1.80 | 0.08 | 2.37 | 1.50 | 0.14 |
| ⟨x⟩ | 22.84 | 18.16 | 0.73 | 2.04 | 0.41 | 1.88 | 5.11 | 0.27 |
| σ | 8.54 | 8.82 | 0.16 | 0.65 | 0.23 | 0.32 | 1.77 | 0.15 |
| min(x) | 5.00 | 3.00 | 0.00 | 1.00 | 0.00 | 1.00 | 1.20 | 0.14 |
| max(x) | 40.00 | 40.00 | 1.00 | 3.44 | 1.00 | 2.37 | 8.74 | 0.90 |

## Genet networks characteristics:

- Clustering larger than random

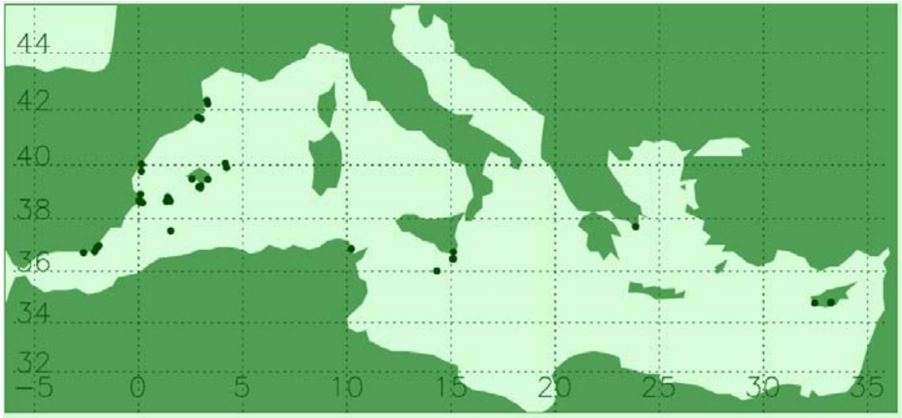
-Diameter of the same order

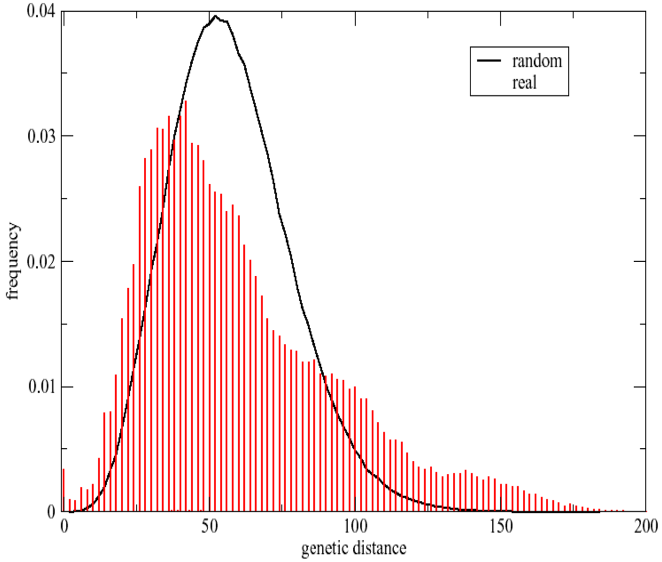
SMALL WORLD

- High density of links

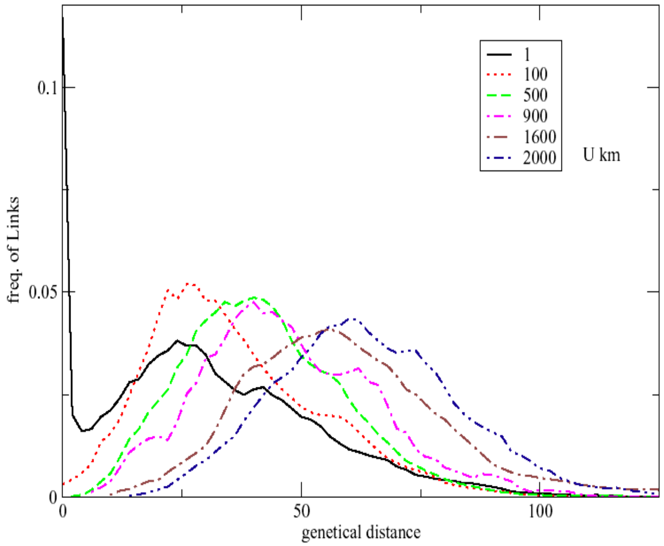
“distance networks”

## Genetic diversity spectrum across the whole Mediterranean



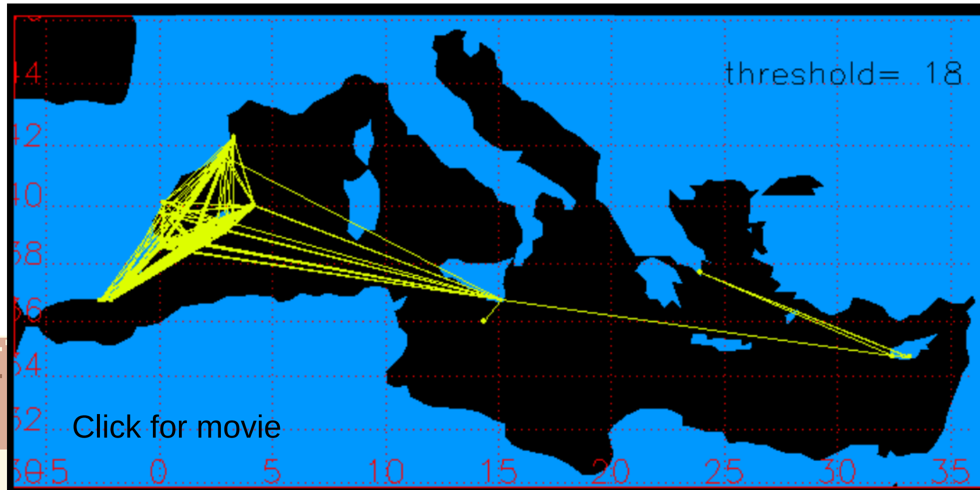



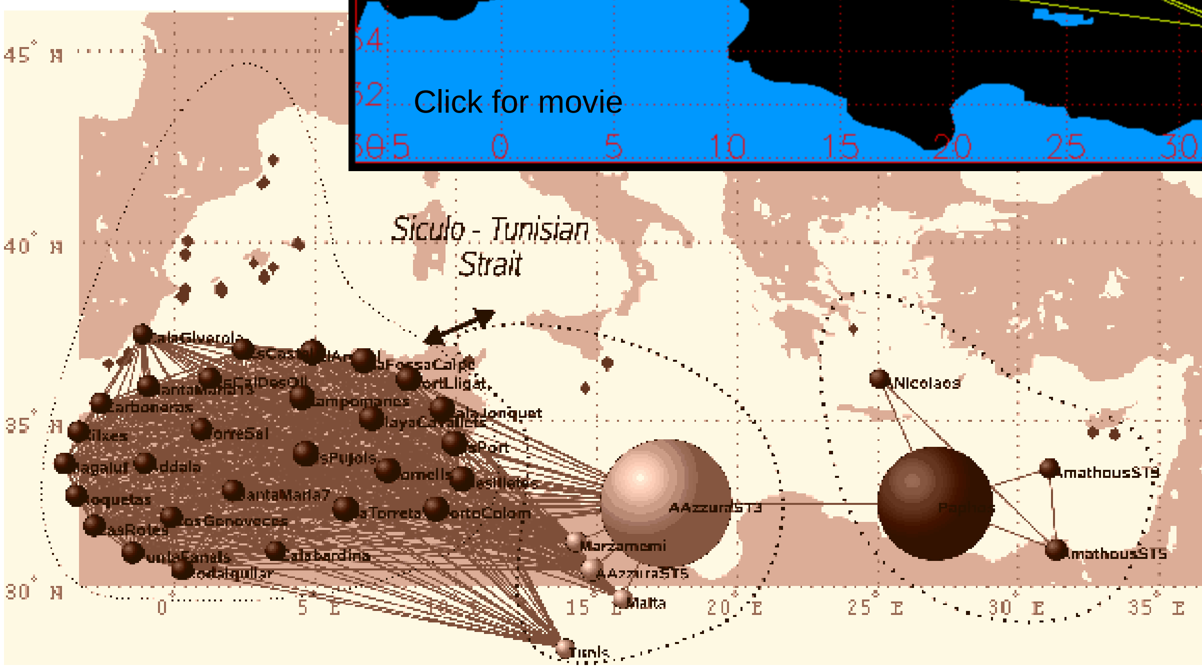

Histogram of Contributions from Links at geographical distance U-100km < d < U

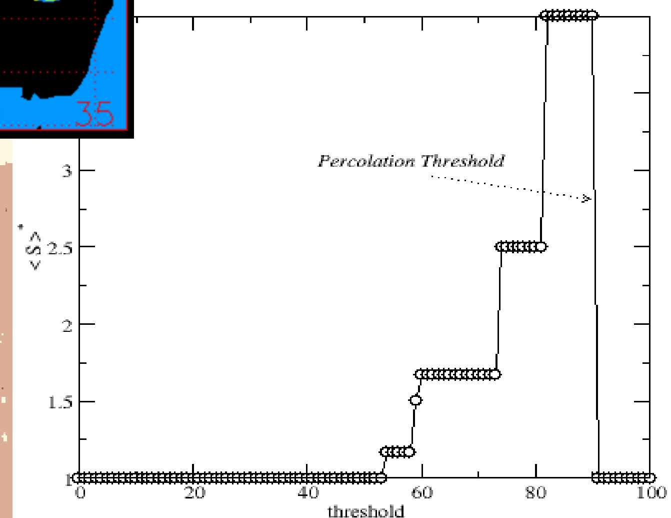


# NETWORKS OF GENETIC SIMILARITY

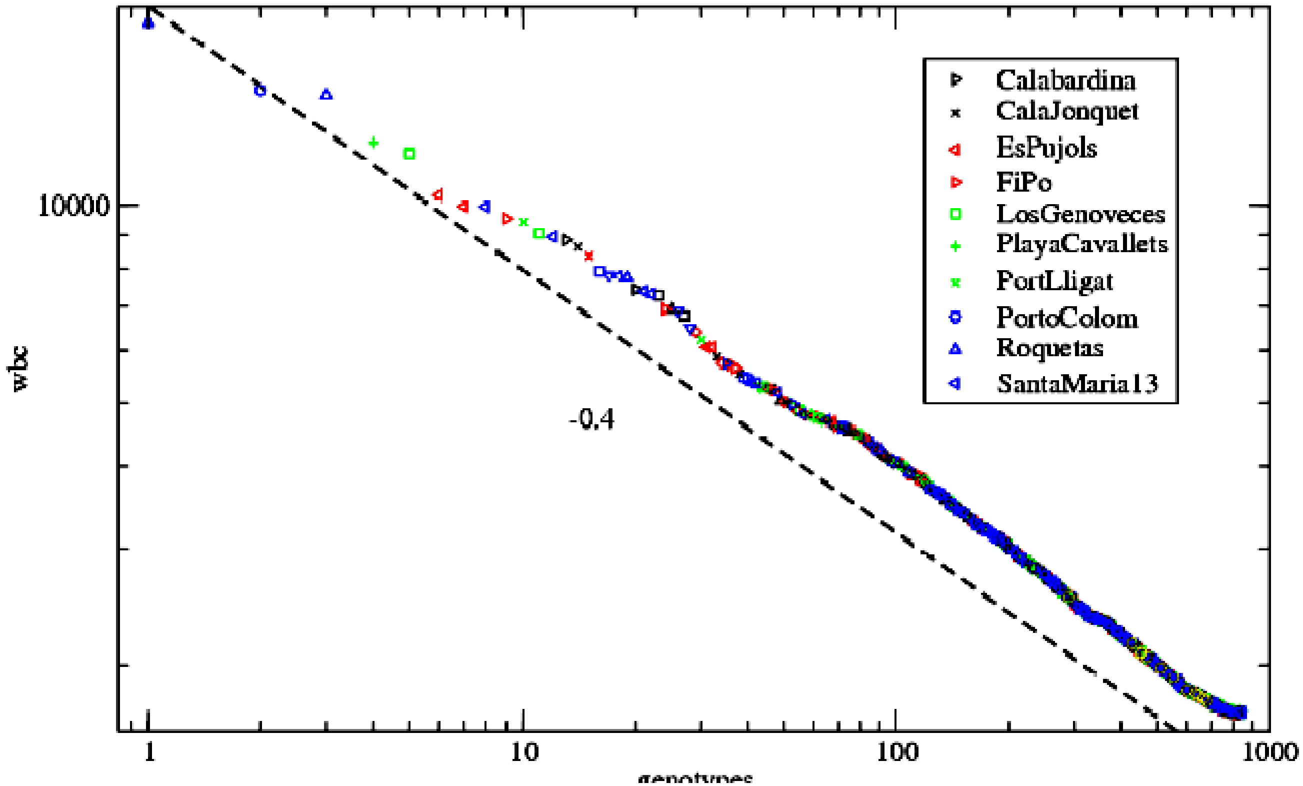
Interpopulation case (metapopulation)



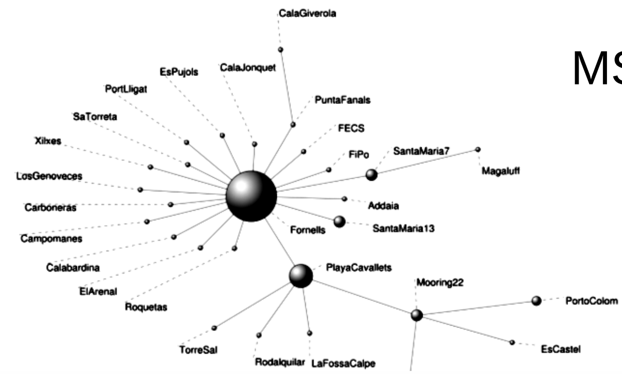
## Genotype betweenness

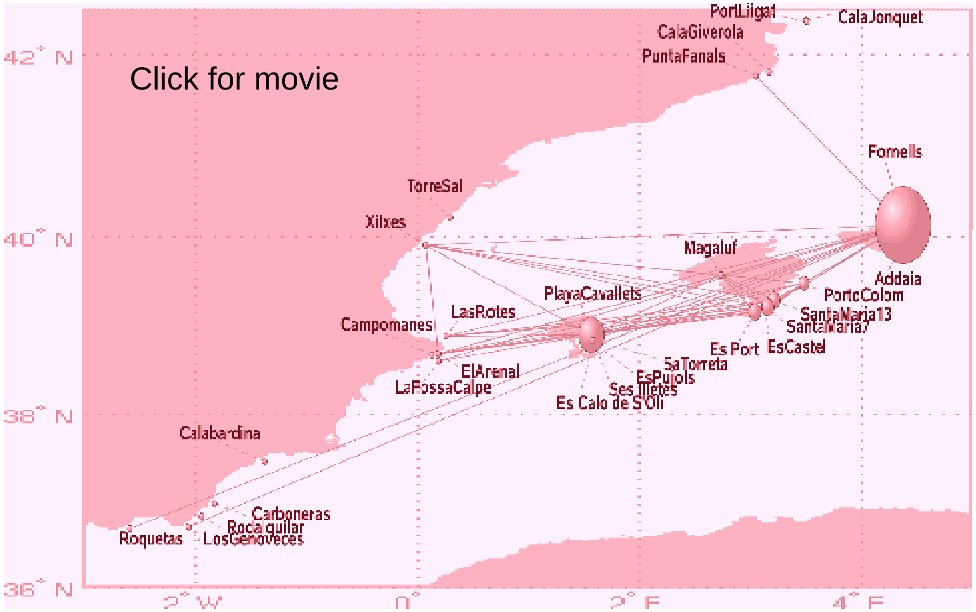



Genotype rank

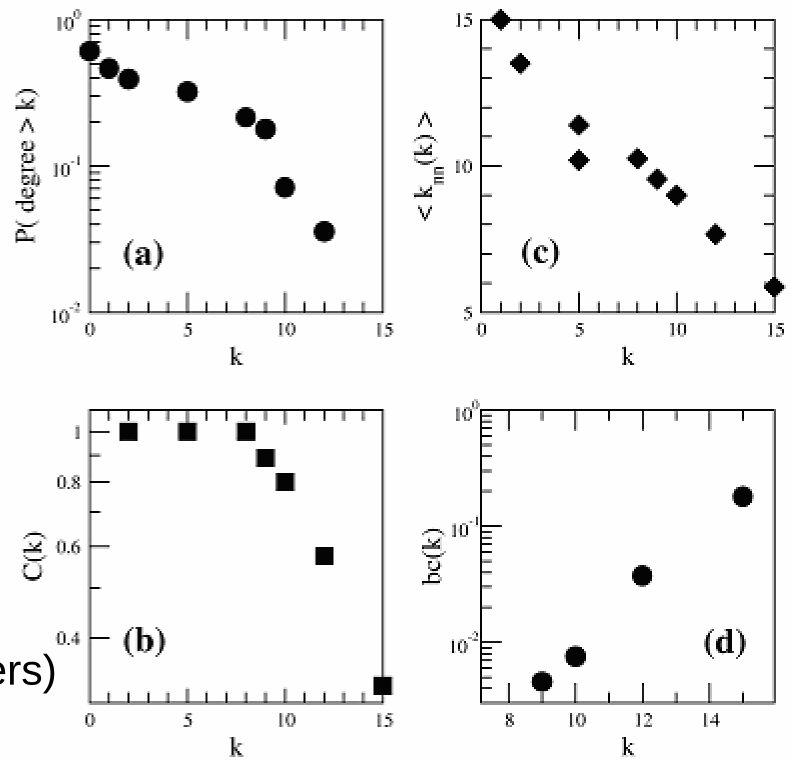
Relevant to gene flow

MST




*dissortativity*




**Population structure:** Nodes in the Balearic Islands have higuest betweeness centrality and degree. “Star-like topology”

**Gene flow:** High clonality and large degree (Fornells and Es Port): net gene sources (exporters)

# SUMMARY

- Intraspecific genetic similarity networks provide powerful tools to summarize genetic information, and give insight into properties such as population structure. The methodology developed is useful for other species beyond Posidonia (Cymodocea is in progress).
- For the organism studied, intrapopulation networks are very dense small worlds. It would be interesting to relate network measures to ecosystem health
- Gene flow, central populations, gene sources, are conveniently revealed by network methods.

Rozenfeld et al., J. Roy. Soc. Interface 4, 1093 (2007)
Rozenfeld et al., submitted (2008)
Hernandez-Garcia et al., AIP Conf. Proc. 913, 78 (2007)
Hernandez-Garcia et al., Physica D 214, 166 (2006)

Available from ifisc.uib.es/publications/
ifisc.uib.es/eng/lines/bio.html

Ifisc.uib.es/EDEN