Network analysis identifies weak and strong links in a metapopulation system

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

1 IMEDEA (CSIC-UIB), Instituto Mediterráneo de Estudios Avanzados, C/ Miquel Marqués 21, 07190 Esporles, Mallorca, Spain.

2 CCMAR, CIMAR-Laboratório Associado, Universidade do Algarve, Gambelas, 8005-139, Faro, Portugal

3 IFISC, Instituto de Física Interdisciplinar y Sistemas Complejos (CSIC-UIB), Campus Universitat de les Illes Balears, E-07122 Palma de Mallorca, Spain.

4 IFREMER, Centre de Brest, BP70, 29280 Plouzané, France

Summary

The identification of key populations for conservation or eradication is a major challenge in population ecology, particularly when dealing threatened, invasive, and pathogenic species. Network theory was applied to map the genetic structure in a metapopulation system using microsatellite data from populations of the threatened seagrass *Posidonia oceanica*, as a model, sampled across its whole geographical range. This approach allowed the characterization of hierarchical population structure, and the identification of populations acting as hubs critical for relaying gene flow and sustaining the metapopulation system. This development opens major perspectives in a broad range of applications of molecular ecology and evolution such as conservation biology and epidemiology, where targeting specific populations is crucial.

Understanding the connectivity between components of a metapopulation system and their role as weak or strong links remains a major challenge of population ecology (*1-3*). Advances in molecular biology fostered the use of indirect approaches to understand metapopulation structure, based on describing the distribution of gene variants (alleles) in space within the theoretical framework of population genetics (*4-7*). Yet, the premises of the classical Wright-Fisher model (*4, 6*) are often violated in real metapopulation systems, such as "migration-drift equilibrium" (*8*), "equal population sizes" or symmetrical rate migration among populations. Threatened or pathogen species, for example, are precisely studied for their state of demographic disequilibrium due to decline and local extinctions in the first case, or to their complex dynamics of local decline and sudden pandemic burst in the second. Furthermore, the underlying hypotheses of equal population size and symmetrical migration rates hamper the identification of putative population "hubs" centralizing migration pathways or acting as sources in a metapopulation system, which is a central issue in conservation biology or epidemiology.

Network theory is emerging as a powerful tool to understand the behavior of complex systems composed of many interacting units (*9-11*). Although network theory has been applied to a broad array of problems (*12-14*), only recently has it been adapted to examining genetic relationships among populations or individuals (*15, 16*). Yet, relevant properties of networks, such as resistance (*9*) to perturbations (i.e. node paralysis or destruction), the ability to host coherent oscillations (*17*) or the predominant importance of nodes or cluster of nodes in maintaining the integrity of the system or relaying information through it can be deducted from the network topology and specific characteristics (*10, 11*). Here we apply network theory to population genetics data of a threatened species, the Mediterranean seagrass *Posidonia oceanica,* to demonstrate its

power to characterize population genetic structure and to identify populations that are critical to the dynamics and sustainability of the whole system. These results open major perspectives in evolutionary ecology, and more specifically in conservation biology and epidemiology where the capacity to target populations deserving major efforts of conservation or control is crucial.

We build networks of population connectivity for a system of 37 meadows of the marine plant *Posidonia oceanica*, sampled across its entire geographic range -the Mediterranean Sea-, by using seven microsatellite markers (*18*). The network was built by considering any pair of populations as linked when their genetic distance (Goldstein distance (*19*)) is smaller than a suitably chosen distance threshold (*20*). We highlight these links as the relevant genetic relationships either at the Mediterranean (the full dataset) or at the regional (28 populations along Spanish coasts) scales.

The topology of the network obtained at the Mediterranean scale (Fig. 1) highlights, without any a priori geographical information being used, the historical cleavage between Eastern and Western basins (*18*) and the transitional position of the populations from the Siculo-Tunisian Strait (see Fig. 1). The average *clustering coefficient* (*20*), $<C>=0.96$, is significantly higher than the one expected after randomly rewiring the links ($<C_0>=0.76$ with $\sigma_0=0.02$, after 10000 randomizations) revealing the existence of clusters of populations more interconnected than expected by chance. The values of *betweenness centrality*, quantifying the relative importance of the meadows in relaying information flow through the network (*20*), immediately highlight a meadow in Sicily (present in 21% of all shortest paths among populations), together with another one in Cyprus (16%), as the main stepping-stones between the pairs of populations sampled in the Western and Eastern basins, respectively (Fig. 1 and Table S1). These

results are in agreement with the genetic structure revealed with classical population genetics analysis (Analysis of Molecular Variance "AMOVA"), revealing past vicariance (*18*) and a secondary contact zone in the Siculo-Tunisian Strait. The metapopulation structure, clustering and 'transition zones' derived from the network analysis arise without any *a priori* input on clustering as needed for AMOVA, and without using geographic information in the analysis of allelic richness previously performed to support the existence and localization of a contact zone *(18)*.

Closer examination of the network conformed by the populations along the Spanish coasts (Figure 2, Table 1), more extensively and homogeneously sampled than the rest of the Mediterranean (Table S1), showed that the *degree distribution*, *P(k*), i.e. the proportion of nodes with *k* connections to other nodes, decays rapidly for large *k* (Fig 3.a) and that the six highest values are all observed in samples collected in the Balearic Islands (Fig 2, Table 1). The average *clustering coefficient* of $<C>$=0.4 was significantly higher than that obtained in the corresponding randomized networks ($<C_0>$=0.13 with σ_0=0.05 after 10000 realizations), whereas the local clustering decays as a function of the degree *k* (Fig 3.b) which indicates that the central core is substructured into a small set of hubs, with high connectivity and low clustering, linking groups of closely connected nodes (i.e. with high clustering). Examination of the relationship between the degree of a node and the average degree of the populations connected to it showed an abundance of links between highly connected and poorly connected nodes (Fig 3.c), a property termed *dissortativity,* present in many biological networks (*21*)**,** and reveals a centralized topology. This observation indicates that seagrass populations along the Spanish continental coasts are genetically closer to Balearic populations than to geographically closer populations. Additionally, the highest values of *betweenness centrality* (Table 1) are also attained at the Balearic populations,

suggesting that the meadows of this region play or have played a central role in relaying gene flow at the scale of the Spanish coasts**.** Moreover, the *betweenness centrality* increases exponentially with the *connectivity degree k* (Fig. 3d). All these findings reveal a star-like structure where hubs are connected in cascade and the central core is the set of Balearic populations. A clear perspective of this pattern is shown by the resulting Minimum Spanning Tree of populations (Fig. S3). The biological implication is a great centrality of the Balearic Islands, acting or having acted as a hub for gene flow thorough the system.

Populations with high degree *k* might either be sources sustaining the system (i.e. spreading propagules), or sinks receiving gene flow from all the other populations, or both. The extremely low rate of sexual recruitment inferred in populations with low clonal diversity (*R*) renders those, if highly connected, much more likely to disperse than to receive. The presence in the Balearic Islands of the two populations with the lowest observed clonal diversity and the highest connectivity (Es Port, *R*=0.1; *k*=10; and Fornells *R*=0.1; *k*=15), likely representing populations supplying "genetic material" to neighbor populations, suggests again that the Balearic islands is a key region for the dynamics and connectivity of the metapopulation system at the scale of the Spanish coast. Furthermore, 8 among 16 continental populations show extreme low connectivity (*k*=0), thereby allowing identification of those least likely to be rescued by other populations, once threatened. As in any genetic approach to metapopulation management, the role of currently observed connectivity in future population rescuing is more important if current connectivity is limited by dispersal ability rather than by competitive interactions that could change in the future in decaying populations. Furthermore, given the particular millenary nature of *P. oceanica* clones, current genetic structure is likely to

integrate patterns of gene flow over past centuries, and thus may not reflect present-day dynamics.

Both networks, that at the scale of the whole Mediterranean (Fig. 1) and that for the Spanish coasts (Fig. 2), presented "small world" properties (*22*), i.e. a diameter (L=1.39 and L=1.63 respectively, (*20*)) shorter than expected for random networks ($<L_0>$=1.47 with σ_0=0.01 and $<L_0>$=2.53 with σ_0=0.15 respectively, after 10000 randomizations) whereas their clustering was much higher (see numerical values above), suggesting a highly hierarchical substructure. This provides clear evidence for the appearance of "short-cuts" in gene flow at multiple geographical scales along the history of this species, indicating rare events of large scale dispersal having a significant impact on the genetic composition of populations.

These results demonstrate that network analyses are powerful tools to examine the structure of gene flow across different geographical scales. The use of specific network properties such as the *betweenness centrality* and the *degree distribution* allowed to identify populations relaying gene flow, or as sources supplying the system as well as those less connected, increasing vulnerability to local extinction. The identification of key populations to maintain the gene flow across the species range is essential to guide conservation strategies for this endangered seagrass. In particular, this methodology successfully revealed the existence of an East-West cleavage in the Mediterranean and of a contact/transition zone in the Siculo-Tunisian Straight without any other *a priori* information such as geographic data or expected clustering of populations. Furthermore, network analysis tools provided graphical representations of the genetic relatedness between populations in a multidimensional space (*15*), free of some of the constraints (e.g. a tree-like structure or binary branching) compulsory in classical methods describing population relationships. Addressing gene flow using network theory may

prove a ground-breaking milestone in critical areas such as conservation biology, dealing with threatened or invasive species, and epidemiology, where the definition of target populations to be conserved or eradicated is of crucial importance.

References

1. D. T. Haydon, S. Cleaveland, L. H. Taylor, M. K. Laurenson, *Emerging Infectious Diseases* **8**, 1468 (2002).
2. J. M. J. Travis, K. J. Park, *Animal Conservation* **7**, 321 (2004).
3. J. C. Williams, C. S. ReVelle, S. A. Levin, *Frontiers in Ecology and the Environment* **2**, 98 (2004).
4. R. A. Fisher, *The genetical theory of natural selection* (Clarendon press, 1930).
5. J. Haldane, *The Causes of Evolution* (Longmans Green, 1932).
6. S. Wright, *Genetics* **16**, 97 (1931).
7. S. Wright, *Genetics* **28**, 114 (1943).
8. J. Hey, C. A. Machado, *Nat. Rev. Genet.* **4**, 535 (2003).
9. R. Albert, H. Jeong, A. L. Barabasi, *Nature* **406**, 378 (2000).
10. L. A. N. Amaral, A. Scala, M. Barthelemy, H. E. Stanley, *Proc. Natl. Acad. Sci. U. S. A.* **97**, 11149 (2000).
11. S. H. Strogatz, *Nature* **410**, 268 (2001).
12. R. M. May, *Trends Ecol. Evol.* **21**, 394 (2006).
13. S. R. Proulx, D. E. L. Promislow, P. C. Phillips, *Trends Ecol. Evol.* **20**, 345 (2005).
14. D. J. Watts, *American Journal of Sociology* **105**, 493 (1999).
15. R. J. Dyer, J. D. Nason, *Mol. Ecol.* **13**, 1713 (2004).
16. A. F. Rozenfeld *et al.*, *Journal of the Royal Society Interface* **4**, 1093 (2007).
17. L. F. Lago-Fernandez, R. Huerta, F. Corbacho, J. A. Siguenza, *Phys. Rev. Lett.* **84**, 2758 (2000).
18. S. Arnaud-Haond *et al.*, *Journal of Biogeography* **34**, 963 (2007).

19. D. B. Goldstein, A. R. Linares, L. L. Cavalli-Sforza, M. W. Feldman, *Proc. Natl. Acad. Sci. U. S. A*. **92**, 6723 (1995).

20. Description of the Methods is available in Science Online as Supplementary Online Material

21. M. E. J. Newman, *Phys. Rev. Lett.* **89**, 208701 (2002).

22. D. J. Watts, S. H. Strogatz, *Nature* **393**, 440 (1998).

23. We acknowledge financial support from the Spanish MEC (Spain) and FEDER through project FISICOS (FIS2007-60327), the Portuguese FCT and FEDER through project NETWORK(POCI/MAR/57342/2004), the BBVA Foundation (Spain), and the European Commission through the NEST-Complexity project EDEN (043251).

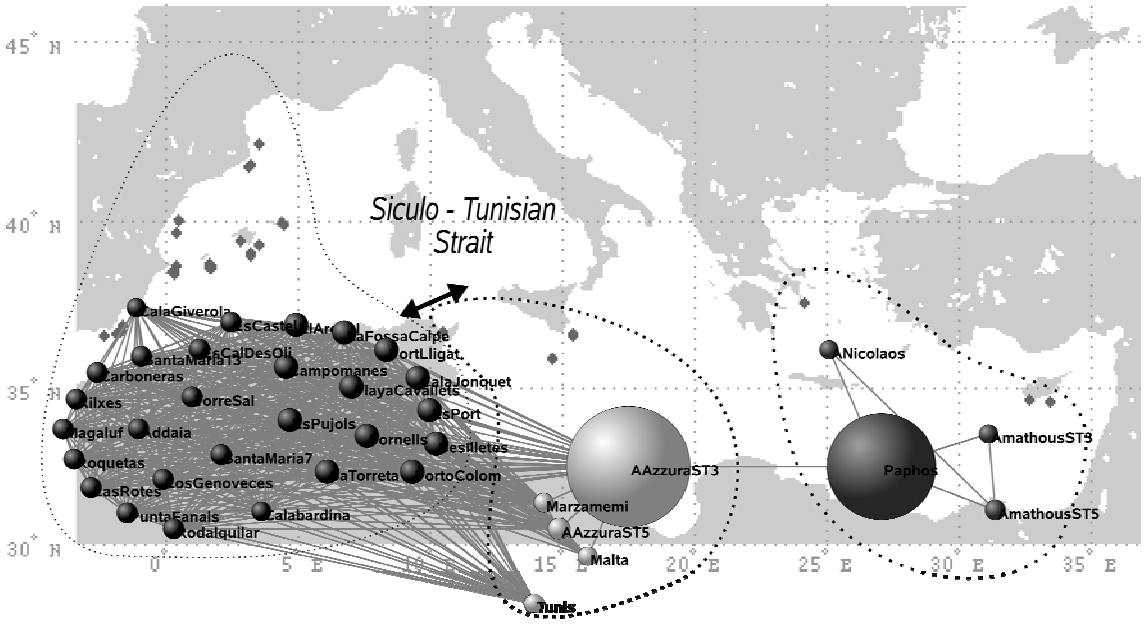

Figure 1: The network of Mediterranean meadows in which only links with Goldstein distances smaller than the percolation distance *Dp*=91 (see Fig. S4) are present (*20*). Nodes representing populations are roughly arranged according to their geographic origin. The precise geographic locations are indicated as diamonds in the background map. One can identify two clusters of meadows, corresponding to the Mediterranean basins (east and west), separated by the Siculo-Tunisian Strait. The size of each node indicates its *betweenness centrality* (*i.e.* the proportion of all shortests paths getting through the node).

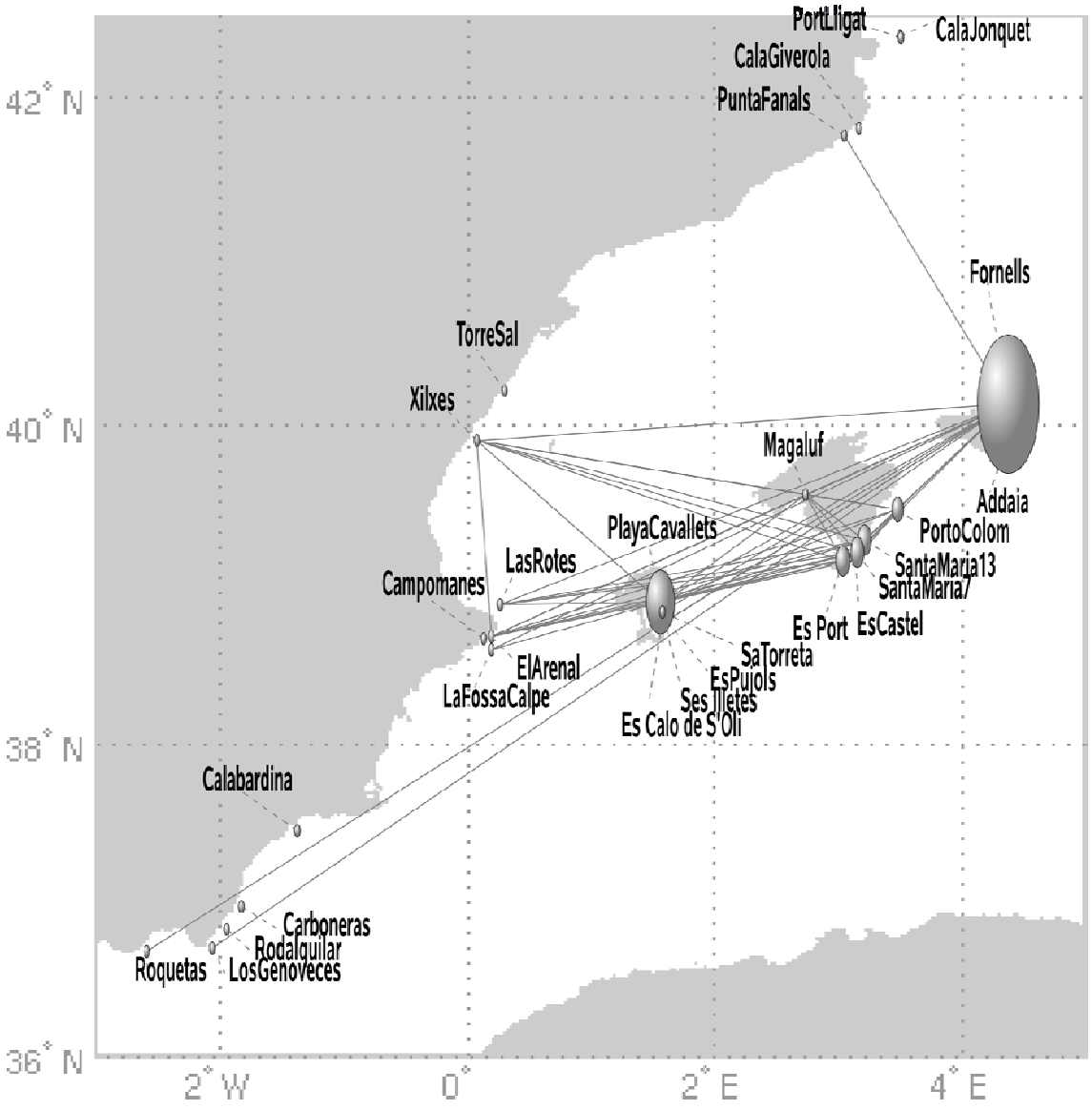

Figure 2: The network constructed with the Spanish meadows (*20*). Nodes are shown at the populations geographic locations. Node sizes characterize their *betweenness centrality* (*i.e.* the proportion of all shortests paths getting through the node).

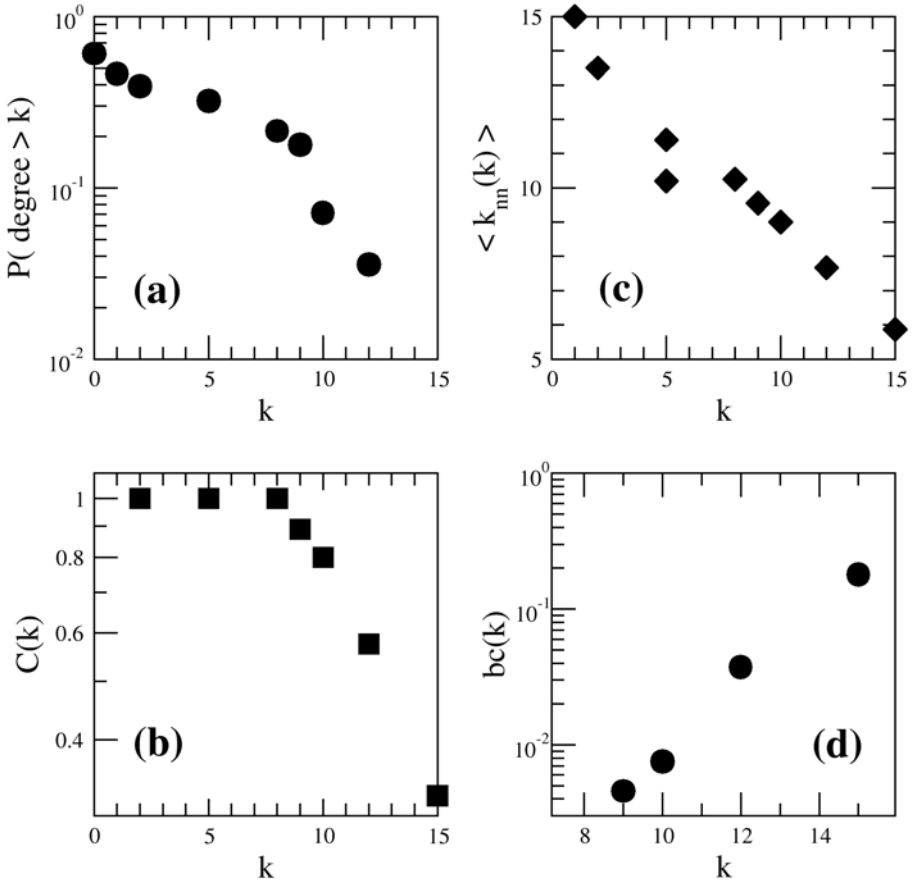

Figure 3: Main topological properties found by analysing the structure of the network of meadows at the Spanish basin scale (Fig. 2). (a) The complementary cumulative *degree distribution* P(degree>k), (b) the local *clustering C(k)*, (c) the average degree $<k_{nn}(k)>$ in the neighbourhood of a meadow with degree *k*, and (d) the *degree-dependent betweenness, bc(k),* as a function of the connectivity degree *k*.

Table 1: Local properties of the network constructed with the Spanish meadows. Information is given for the *connectivity degree* (*k*), *betweenness centrality* (bc) and *clustering* (*C*), as well as clonal diversity estimates (*R*) for each sample.

REGION		Name	R	k	bc	C
SPANISH BALEARIC ISLANDS	*Menorca*	Addaia	0,67	0	0	0
		Fornells	0,1	15	0,180	0,32
	Mallorca	Magaluf	0,68	8	0	1
		Porto Colom	0,5	9	0,0046	0,89
	Cabrera	Es Castel	0,1	5	0	1
		Es Port	0,1	10	0,0075	0,8
		Santa Maria 13	0,56	10	0,0075	0,8
		Santa Maria 7	0,54	10	0,0075	0,8
	Ibiza	Playa Cavallets	0,73	12	0,0037	0,58
		Es Pujols	0,67	1	0	0
	Formentera	EsCalo de S'Oli	0,36	0	0	0
		Ses Illetes	0,6	0	0	0
		Sa Torreta	0,51	2	0	1

REGION		Name	R	k	bc	C
SPANISH IBERIAN PENINSULA	*(ORDERED FROM NORTH TO SOUTH)*	Cala Jonquet	0,5	0	0	0
		Port Lligat	0,28	0	0	0
		Cala Giverola	0,43	0	0	0
		Punta Fanals	0,68	1	0	0
		Torre Sal	0,5	0	0	0
		Xilxes	0,35	8	0	1
		Las Rotes	0,73	5	0	1
		El Arenal	0,86	8	0	1
		Campomanes	0,7	0	0	0
		LaFossaCalpe	0,77	2	0	1
		Calabardina	0,88	0	0	0
		Carboneras	0,32	0	0	0
		Rodalquilar	0,53	0	0	0
		Los Genoveces	0,34	1	0	0
		Roquetas	0,69	1	0	0

Supporting Online Material

METHODS

Molecular data. About 40 *Posidonia* shoots collected at each of the 37 sampled populations (Fig.1 and Table 2) were genotyped with a previously selected set of seven dinucleotide microsatellites (*1*) allowing the identification of clones (also called genets for clonal plants). Clonal diversity was estimated for each population as described in Arnaud-Haond et al. (*1*), and replicates of the same clone were excluded for the estimation of inter-population distances. The matrix of interpopulation distances was built using Goldstein metrics (*2*), thus taking into account the level of molecular divergence among alleles, besides the differences in terms of allelic frequencies.

Networks. We first built a fully connected network with the 37 populations considered as nodes. Each link joining pairs of populations was labeled with the Goldstein distance among them. We then removed links from this network of genetic similarity, starting from the one with the largest genetic distance and following in decreasing order, until the network reaches the percolation point (*3, 4*), beyond which it loses its integrity and fragments into small clusters. This means that gene flow across the whole system is disabled if connections at a distance smaller than this critical one, *Dp*, are removed. The precise location of this percolation point is made with the standard methodology adequate for finite systems (*3, 4*), i.e., by calculating the average size of the clusters excluding the largest one, $\langle S \rangle^{*} = \frac{1}{N} \sum_{s<S_{\max}} s^2 n_s$, as a function of the last distance value removed, *thr*, and identifying the critical distance with the one at which <S>* has a maximum. ***N*** is the total number of nodes not included in the largest cluster

and n_s is the number of clusters containing s nodes. Here we find Dp=91, as shown in Fig. S1.

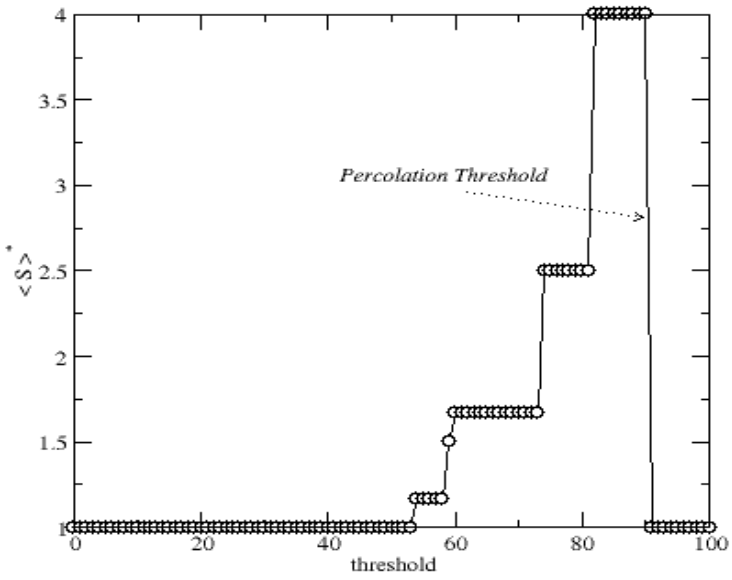

Figure S1. The average cluster size excluding the largest one, as a function of the imposed genetic threshold, at the whole Mediterranean scale. This identifies Dp=91 as the percolation threshold.

Once the network at percolation point is obtained, we analyzed its topology and characteristics (See Table 2 below, and Fig.1), and interpret those biologically. The first column in Table 2 contains also the estimated clonal diversity R of the different populations, defined as the proportion of different genotypes found with respect to the total number of collected shoots.

Table S1: Local properties of the whole Mediterranean network for *thr*=*Dp*=91. Information is given for the *betweenness centrality* (bc) and *clustering* (*C*), as well as clonal diversity estimates (*R*) for each sample.

REGION		Name	R	bc	C
SPANISH BALEARIC ISLANDS	*Menorca*	Addaia	0,67	0,0010	0,980
		Fornells	0,1	0,0031	0,946
	Mallorca	Magaluf	0,68	0,0010	0,981
		Porto Colom	0,5	0,0031	0,946
		Es Castel	0,1	0,0010	0,981
	Cabrera	Es Port	0,1	0,0031	0,946
		Santa Maria 13	0,56	0,0010	0,981
		Santa Maria 7	0,54	0,0010	0,981
	Ibiza	Playa Cavallets	0,73	0,0031	0,946
		Es Pujols	0,67	0,0031	0,946
	Formentera	EsCalo de S'Oli	0,36	0,0010	0,981
		Ses Illetes	0,6	0,0031	0,946
		Sa Torreta	0,51	0,0031	0,946
CENTRAL BASIN	*Sicily*	Tunis	0,85	0	1
		Malta	0,74	0	1
		A. AzzuraST3	0,77	0,205	0,897
		A. AzzuraST5	0,72	0,0017	0,963
		Marzamemi	0,81	0,0003	0,995

REGION		Name	R	bc	C
SPANISH IBERIAN PENINSULA	*(ORDERED FROM NORTH TO SOUTH)*	Cala Jonquet	0,5	0,0031	0,946
		Port Lligat	0,28	0,0031	0,946
		Cala Giverola	0,43	0	0,997
		Punta Fanals	0,68	0,0010	0,981
		Torre Sal	0,5	0,0010	0,981
		Xilxes	0,35	0,0010	0,981
		Las Rotes	0,73	0,0010	0,981
		El Arenal	0,86	0,0031	0,946
		Campomanes	0,7	0,0031	0,946
		LaFossaCalpe	0,77	0,0031	0,946
		Calabardina	0,88	0,0003	0,997
		Carboneras	0,32	0,0010	0,981
		Rodalquilar	0,53	0,0010	0,981
		Los Genoveces	0,34	0,0010	0,981
		Roquetas	0,69	0,0010	0,981
EAST BASIN	*Cyprus*	Amathous ST3	0,44	0	1
		Amathous ST5	0,62	0,0008	0,667
		Paphos	0,68	0,1579	0,333
	Greece	A. Nicolaos	0,69	0	1

At the Spanish coasts scale, no percolation point is found using the above procedure, meaning that the genetic structure in this area is rather different from the one at the whole Mediterranean scale. To construct a useful network representation of the meadows genetic similarity, the following alternative process was applied in order to determine a relevant distance threshold, *thr*, above which links are discarded. At a very low threshold (*thr*=16, see Movie S1) only the inner part of a central core, constituted by some meadows from the Balearic Islands, is connected. As the threshold is increased new meadows (from the central Spanish coast) become connected (*thr*=20). Beyond that value, more peripheral meadows are connected from the northern and southern Spanish coasts. The geographical extension of the connected cluster (Fig. S2) grows with the distance threshold and an important jump occurs at *thr*=22, when the northern and southern coasts get connected for the first time.

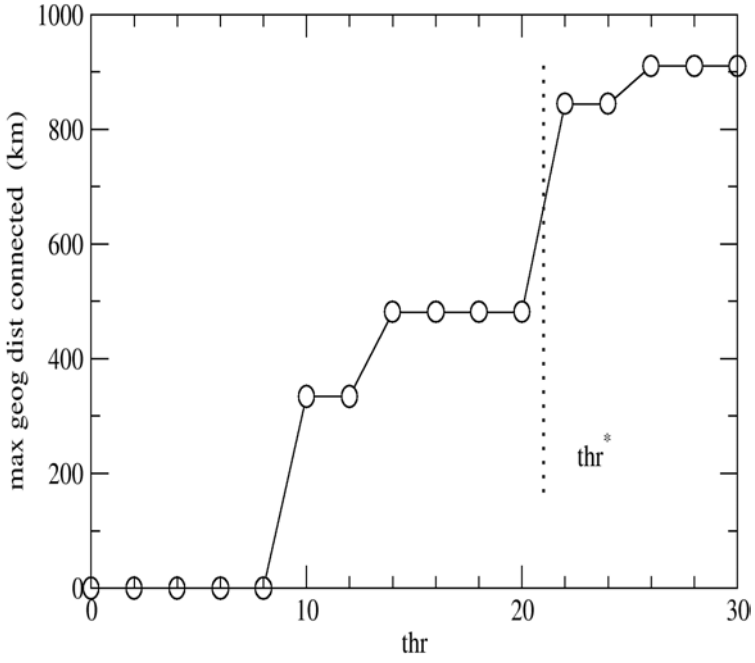

Figure S2. The maximal geographic distance connected (at the Spanish coasts scale) as a function of the imposed distance threshold (*thr)*. Above *thr*=22 the maximal geographic distance covered by connected populations nearly duplicates.

Further distance threshold increase does not contribute to geographical extension. Therefore, we find the value *thr*=22 and the resulting network as appropriate for topological characterization, since at this point the network contains a rich mixture of strong and weak links spanning all the available geographic scales within the Mediterranean Spanish coasts.

Estimates of global and local properties of the network. The degree k_i of a given node i is the number of other nodes linked to it (i.e., the number of neighbor nodes). The *distribution* $P(k)$ gives the proportion of nodes in the network having degree k.

We denote by E_i the number of links existing among the neighbors of node i. This quantity takes values between 0 and $E_i^{(\max)} = \frac{k_i(k_i-1)}{2}$, which is the case of a fully connected neighborhood. The clustering coefficient C_i of node i is defined as:

$$C_i = \frac{E_i}{E_i^{(\max)}} = \frac{2E_i}{k_i(k_i-1)}$$

The clustering coefficient of the whole network $<C>$ is defined as the average of all individual clustering coefficients in the system. The degree dependent clustering $C(k)$ is obtained after averaging C_i for nodes with degree k.

Real networks exhibit correlations among their nodes (*5-11*) that play an important role in the characterization of the network topology. Those node correlations are furthermore essential to understand the dynamical aspects such as spreading of information or their robustness against targeted or random removal of their elements. In social networks, nodes having many connections tend to be connected with other highly connected nodes. This characteristic is usually referred to as *assortativity*, or *assortative mixing*. On the other hand, technological and biological networks show rather the

property that nodes having high degrees are preferably connected with nodes having low degrees, a property referred to as *dissortativity*. Assortativity is usually studied by determining the properties of the average degree $<k_{nn}>$ of neighbors of a node as a function of its degree *k*. (*5, 6, 12*)**.** If this function is increasing, the network is assortative, since it shows that nodes of high degree connect, on average, to nodes of high degree. Alternatively, if the function is decreasing, the network is dissortative, as nodes of high degree tend to connect to nodes of lower degree. In this last case, the nodes with high degree are therefore central hubs ensuring the connection of the whole system.

The betweenness centrality of node *i*, *bc(i*), (*13*) counts the fraction of shortest paths between pairs of nodes which pass through node *i*. Let σ_{st} denote the number of shortest paths connecting nodes *s* and *t* and $\sigma_{st}(i)$ the number of those passing through the node *i*. Then,

$$bc(i) = \sum_{s \neq t \neq i} \frac{\sigma_{st}(i)}{\sigma_{st}}.$$

The degree-dependent betweeness, *bc(k)*, is the average betweeness value of nodes having degree *k*.

Minimum Spanning Tree.

Given a connected, undirected graph, a spanning tree of that graph is a subgraph without cycles which connects all the vertices together. A single graph can have many different spanning trees. Provided each edge is labeled with a cost (in our analysis the genetic distance among the connected populations) each spanning tree can be characterized by the sum of the cost of its edges. A minimum spanning tree is then a spanning tree with

minimal total cost. A minimum spanning tree is in fact the minimum-cost subgraph connecting all vertices, since subgraphs containing cycles necessarily have more total cost. Figure S3 shows the minimum spanning tree for the Spanish meadows. The star-like structure centered at Balearic populations is evident.

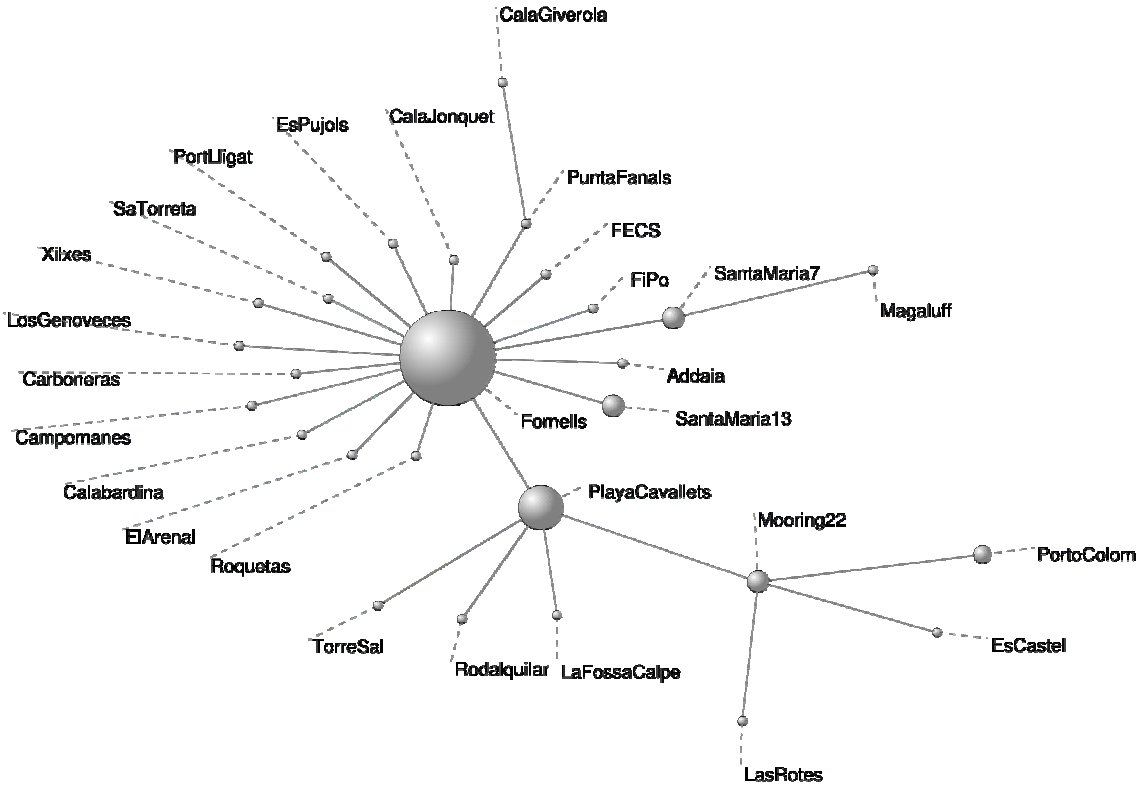

Figure S3: Minimum Spanning Tree based on Goldstein distance among Spanish meadows. This is the subgraph which connects the populations at the Spanish coast scale minimizing the total genetic distance along links.

Movie S1

Sequence of networks for the Spanish populations obtained at successive values of the threshold distance, *thr*, above which links are discarded.

Supporting References

1. S. Arnaud-Haond *et al.*, *J. Hered.* **96**, 434 (2005).

2. D. B. Goldstein, A. R. Linares, L. L. Cavalli-Sforza, M. W. Feldman, *Proc. Natl. Acad. Sci. U. S. A.* **92**, 6723 (1995).

3. G. Grimmett, *Percolation*, A Series of Comprehensive Studies in Mathematics (Springer-Verlag, Berlin, ed. 2nd edition, 1999).

4. D. Stauffer, A. Aharony, *Introduction to Percolation Theory* (Taylor & Francis, CRC Press, London, ed. 2nd. edition, 1994).

5. R. Pastor-Satorras, A. Vazquez, A. Vespignani, *Phys. Rev. Lett.* **87**, 258701 (2001).

6. M. E. J. Newman, *Phys. Rev. Lett.* **89**, 208701 (2002).

7. S. Maslov, K. Sneppen, *Science* **296**, 910 (2002).

8. A. Vazquez, R. Pastor-Satorras, A. Vespignani, *Phys. Rev. E* **65**, 066130 (2002).

9. M. E. J. Newman, *Phys. Rev. E* **67**, 026126 (2003).

10. M. Boguñá, R. Pastor-Satorras, *Phys. Rev. E* **68**, 036112 (2003).

11. A. Barrat, R. Pastor-Satorras, *Phys. Rev. E* **71**, 036127 (2005).

12. S. H. Lee, P. J. Kim, H. Jeong, *Phys. Rev. E* **73**, 016102 (2006).

13. L. C. Freeman, *Sociometry* **40**, 35 (1977).