PALE-Blu work package 4, deliverable report (M54)

D4.5: Description of ecological factors driving population structure and of natural barriers to vector dispersion (Partner 3, 6, 7, 12)

Background and deliverable description

One of the objective of the WP4 is to compare distribution, structure, evolutionary history and connectivity of *Culicoides* populations and the different bluetongue virus (BTV) epi-zones in Europe defined in WP3. Indeed, understanding how geographical and environmental features, landscape patterns and in particular those created by human land uses, structure vector populations and constraint vector dispersion will help to understand factors driven the definition of epizones (D3.3). *Culicoides* dispersal capacities are described as 'dispersive stratified', which results from the combination of processes occurring at a short distance combined with jumps over long distances. Although winds and local air movements have an undeniable role in the dispersal of *Culicoides*, anthropogenic factors also play a predominant role, particularly at the local level (host distribution, landscape patterns).

Several studies using direct measure of *Culicoides* dispersal (mark-release-recapture studies) have been carried out, but methodological constraints, such as extremely low recapture rate or as doubts about similar dispersal capacities between wild *Culicoides* and captured-marked *Culicoides*, limit the interpretations of results. Thus, indirect methods using population genetics have been successfully applied to *Culicoides* to assess the local range expansion of *C. imicola* in continental France (Jacquet *et al.*, 2016a) and to reconstruct the chain of events that led to a newly colonized French area nestled at the northern foot of the Pyrenees by *C. imicola* (Jacquet *et al.*, 2016b).

The D4.5 activities aimed to determine the inland connectivity of populations of *C. obsoletus s.s.* at large geographical scales and to identify the environmental factors that promote or limit gene flow between *Culicoides* populations in France, using an innovative framework integrating spatial, population genetics and statistical approaches. Therefore, the D4.5 activities determine, by observing patterns of allelic frequency distributions, how dispersion and historic events may have shaped the spatial arrangement of genetic diversity in Western Europe.

Publication related to this deliverable

 Mignotte A, Garros C, Dellicour S, Jacquot M, Gilbert M, Gardès L, Balenghien T, Duhayon M, Rakotoarivony I, de Wavrechin M, Huber K (2021). High dispersal capacity of *Culicoides obsoletus* (Diptera: Ceratopogonidae), vector of bluetongue and Schmallenberg viruses, revealed by landscape genetic analyses. Parasites & Vectors 14, 93. doi: 10.1186/s13071-020-04522-3.

Deliverable Activities

1. Important dispersal capabilities of Culicoides obsoletus s.s.

During an epidemic of a vector-borne disease, a precise understanding of the dispersal capacity of vector species is a key factor in implementing appropriate control strategies to limit the spread of the disease. *Culicoides* dispersal is described as stratified, due to a combination of dispersal processes occurring actively at short distances, and passively or semi-actively at long distances (Murray and Kirkland 1995). This dispersion is dependent on landscape characteristics ("environmental factors") which will favor (conductance factor) or limit (resistance factor) the dispersal.

Partner 3 aimed to determine the inland connectivity of populations of *C. obsoletus s.s.* at large geographical scales and to identify the environmental factors that promote or limit gene flow between *Culicoides* populations in France, using an innovative framework integrating spatial, population genetics and statistical approaches (Figure 1). Partner 3 used an individual-based approach that considers that *Culicoides* populations are probably not spatially structured into separate populations and must be considered as a continuum of individuals heterogeneously distributed across a landscape. This approach in landscape genetics aims at maximizing the number of sampling sites, and thus brings much greater statistical power to the detection of spatial patterns of genetic differentiation and the environmental factors that cause them (Landguth et al. 2010). A total of 348 individuals were sampled in 46 sites in France and were genotyped using 13 newly designed microsatellite markers.



Figure 1. Analytical workflow for population genetics and landscape analyses conducted in the context of the study of environmental factors influencing gene flow in *C. obsoletus s.s.* at the French scale.

Low genetic differentiation and a weak population structure for *C. obsoletus s.s.* across the country were found. Using three complementary inter-individual genetic distances, Partner 3 did not detect any significant isolation by distance (Figure 2). This reflects a high level of gene flow and reinforces what has been described previously for other *Culicoides* species such as *C. imicola* in France (Jacquet *et al.* 2016a), *C. brevitarsis* in Australia (Onyango *et al.* 2015) or *C. stellifer* in North America (Shults *et al.* 2020).



Figure 2. Genetic clustering and genetic differentiation of *Culicoides obsoletus s.s.* Results of the genetic clustering analyses performed with GENELAND (**a**) and STRUCTURE (**b**), as well as smoothing of pairwise measures performed with mapping averaged pairwise information (MAPI) and based on (**c**) Rousset's (a_R) and (**d**) factorial correspondence analysis (FCA) inter-individual genetic distances. A specific colour has been assigned to each genetic cluster in **a** and **b**. **c**, **d**. Genetic dissimilarity is represented by a colour scale ranging from red (lower genetic dissimilarity) to blue (higher genetic dissimilarity). The black circles indicate the sampling sites.

Partner 3 employed a multiple regression on distance matrices approach to investigate the correlation between genetic and environmental distances. While the unique contribution of the environmental factors tested was very small, the contribution of cattle density was the strongest that we detected. Thus, active host-seeking movements of *Culicoides* may have contributed to homogenize the genetic diversity. These results are in line with those obtained in landscape genetics studies of BTV, which identified distributions of cattle and sheep as key factors in BTV dispersal (Jacquot *et al.* 2017).

The most genetically dissimilar individuals were mainly from the southernmost populations of the sampling area. Wind-favored dispersal of *Culicoides* could explain the significant anisotropic isolation by distance (Jay *et al.* 2012), i.e., direction-dependent, observed along the north/south axis in France (Figure 3).



Figure 3. Anisotropic isolation by distance, correlation between geographic distances projected by angle and interindividual genetic distances (a_R , (Rousset 2000)). The projected distance matrix was calculated for each angle between all sampling sites, using the formula presented. A linear regression of the genetic distances on the projected geographical distances obtained for each angle was then performed. The angle that maximizes the R² of this regression with a positive regression coefficient was considered as the angle maximizing the isolation by distance signal. (B) Bearing analysis, correlation between genetic (a_R) and geographical distances as a function of the angle between sampling sites.

Dispersion phenomena caused by wind currents, mainly over the seas, have already been established for *Culicoides* (Murray and Kirkland 1995, Alba *et al.* 2004, Gloster *et al.* 2008, Jacquet *et al.* 2016b). The resultant dispersal events can be described as passive and active, respectively, because recapture was achieved downwind and upwind of the prevailing wind direction (Kluiters *et al.* 2015). Examination of the map of the average wind directions in France over the last 10 years provides potential explanations for this (Figure 4). It can be seen that the southernmost sampling sites (with the most dissimilar individuals) are located in an area where the wind direction is different from that in the rest of France. It should be noted that the diffusion of BTV, and thus the dispersion of *Culicoides*, has already been associated with wind direction (Sedda *et al.* 2012, Kluiters *et al.* 2015, Sedda *et al.* 2015).



Figure 4. Map of wind direction averaged from 2000 to 2010. Sampling sites are represented by black points. The colour scale represents the wind direction from 0 to 360 ° from north. The arrowheads indicate the exact wind direction of each raster pixel.

This study shows that the genetic structure of populations at the scale of a country can become homogeneous through large-scale dispersion. These results demonstrate, for *C. obsoletus s.s.*, a very high inland dispersal and vectorization capacity, which has to be taken into consideration in further work on vector competence and epidemiological modeling of disease transmission.

2. Analysis of the diversity and genetic structure of C. obsoletus s.s. in Europe

The objective of this part was to determine, by observing patterns of allelic frequency distributions, how dispersion and historic events may have shaped the spatial arrangement of genetic diversity. This study was conducted at continental scale, on two transects from Morocco to Finland and from Turkey to the United Kingdom. Partner 3 used 11 microsatellite markers and a mitochondrial gene (Cox1), to identify genetic and phylogeographic structure of *C. obsoletus s.s.* at the European scale and determine historical events that have shaped the current genetic structure pattern.

This study revealed low levels of genetic differentiation between populations at the European scale, with significant gene flow at the continental level and high connectivity between populations separated by large geographic distances. Bayesian clustering methods reveal four genetic clusters distributed all over the study area but one of the clusters is mainly represented in the Balkan area and another one encompassed the Canary Islands (Figure 5).



Figure 5. Microsatellite genetic clustering and genetic diversity of *C. obsoletus s.s.* sampled populations. Results of the genetic clustering analyses performed with STRUCTURE (a) and rarefied allelic richness of populations (b). Panels (1) corresponds to an enlargement of the geographical area (a) of the panel circled in red. One specific color has been assigned to each genetic cluster in panels (a) and (1).

A significant but low level of isolation by distance in continental Europe was observed suggesting a steppingstone model of migration compatible with the combination of passive and active dispersal capacity of *Culicoides* species as it has already been demonstrated for *C. imicola* (Jacquet *et al.* 2016b). The passive dispersion of the *Culicoides* by winds over long distances certainly contributes to blurring the signal of isolation by distance. Only the population of the Canary Islands is well genetically differentiated. This population also exhibits low allelic richness and a very low nucleotide diversity. Island species are likely to have experienced population bottlenecks at some point in their evolutionary history due to founder events during the initial island colonization.

The haplotype mismatch distribution, the lack of matrilineal subdivision, the star-like pattern of medianjoining haplotype network centered around few frequent haplotypes (Figure 6) combined to significant negative values of Fu's (Fu 1997) and Tajima's D (Tajima 1989) suggest a rapid demographic expansion of *C. obsoletus s.s.* in Europe. Recent demographic expansion is consistent with demographic history of most European species modified during the glacial episodes of the Pleistocene (Hewitt 2004, Phillips *et al.* 2010). Given the dependence of *Culicoides* on bovine hosts, the development of modern human agriculture may have played an important role in the population explosion of *C. obsoletus s.s.* in Europe.



Figure 6. Median-joining haplotype networks of cox1 mitochondrial sequences of *C. obsoletus s.s.* The size of the circles is proportional to the number of individuals with that haplotype. Colors represent the country of sampled individuals. The length of the branches separating haplotypes is proportional to the number of mutational steps between them.

The low species diversity within the Obsoletus/Scoticus Complex observed in southeastern Europe and the Balkans (Mignotte *et al.* 2020), combined with the high allelic richness observed for *C. obsoletus s.s.* in this region, led us to hypothesize that this area could have served as a refuge for this species during the Pleistocene. Several scenarios were tested using the ABC method (Figure 7). the most probable scenario identifying the Balkans and the arctic region as the post-glacial origin of ancestral populations, 5,000 to 10,000 years ago, which after more recent admixture events (2,500 years ago) gave rise to the populations of Western Europe (Scenario 1, Figure 7). The DIYABC software used to conduct these analyses does not allow to include recurrent migration phenomena in the tested scenarios (Cornuet *et al.* 2014), which can be a hindrance to study the evolutionary history of species with high dispersal capacities such as *C. obsoletus s.s.* Analyses are currently underway with the ABCtoolbox software (Wegmann *et al.* 2010), which takes migration into account, in order to refine these scenarios.



Figure 7. Graphical representation of the 4 scenarios simulated using an approximate Bayesian computation (ABC) approach on microsatellite data. Red circles correspond to a bottleneck event. T1, T2 and T3 correspond to successive dates of divergence times. NO: population from Norway (Narvik), SP: population from Spain (Torre del campo), FR: population from France (Les Hitres), GR: population from Greece (Velestino).

References

Alba AP, Casal J, Domingo M (2004). Possible introduction of bluetongue into the Balearic Islands, Spain, in 2000, via air streams. Veterinary Record, 155: 460-461.

Cornuet JM, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R, Marin JM, Estoup A (2014). DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. Bioinformatics, 30: 1187-1189.

Fu YX (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics, 147: 915-925.

Gloster J, Burgin L, Witham C, Athanassiadou M, Mellor PS (2008). Bluetongue in the United Kingdom and northern Europe in 2007 and key issues for 2008. Veterinary Record, 162: 298-302.

Hewitt GM (2004). Genetic consequences of climatic oscillations in the Quaternary. Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 359: 183-195.

Jacquet S, Huber K, Guis H, Setier-Rio ML, Goffredo M, Allene X, *et al.* (2016a). Spatio-temporal genetic variation of the biting midge vector species *Culicoides imicola* (Ceratopogonidae) Kieffer in France. Parasites & Vectors, 9: 141. doi: 10.1186/s13071-016-1426-4.

Jacquet S, Huber K, Pagès N, Talavera S, Burgin LE, Carpenter S, *et al.* (2016b). Range expansion of the Bluetongue vector, *Culicoides imicola*, in continental France likely due to rare wind-transport events. Scientific Reports, 6: 27247. doi: 10.1038/srep27247.

Jacquot M, Nomikou K, Palmarini M, Mertens P, Biek R (2017). Bluetongue virus spread in Europe is a consequence of climatic, landscape and vertebrate host factors as revealed by phylogeographic inference. Proceedings. Biological sciences, 284.

Jay F, Sjödin P, Jakobsson M, Blum MGB. 2012. Anisotropic Isolation by Distance: The Main Orientations of Human Genetic Differentiation. Molecular Biology and Evolution, 30: 513-525.

Kluiters G, Swales H, Baylis M (2015). Local dispersal of palaearctic *Culicoides* biting midges estimated by mark-release-recapture. Parasites & Vectors, 8.

Landguth EL, Cushman SA, Murphy MA, Luikart G (2010). Relationships between migration rates and landscape resistance assessed using individual-based simulations. Molecular Ecology Resources, 10: 854-862.

Mignotte A, Garros C, Gardès L, Balenghien T, Duhayon M, Rakotoarivony I, *et al.* (2020). The tree that hides the forest: cryptic diversity and phylogenetic relationships in the Palaearctic vector Obsoletus/Scoticus Complex (Diptera: Ceratopogonidae) at the European level. Parasites & Vectors 13: 265.

Murray M, Kirkland P (1995). Bluetongue and Douglas virus activity in New South Wales in 1989: further evidence for long-distance dispersal of the biting midge *Culicoides brevitarsis*. Australian Veterinary Journal, 72: 56-57.

Onyango MG, Beebe NW, Gopurenko D, Bellis G, Nicholas A, Ogugo M, Djikeng A, Kemp S, Walker PJ, Duchemin JB (2015). Assessment of population genetic structure in the arbovirus vector midge, *Culicoides brevitarsis* (Diptera: Ceratopogonidae), using multi-locus DNA microsatellites. Veterinary Research, 46: 108.

Phillips BL, Brown GP, Shine R (2010). Life-history evolution in range-shifting populations. Ecology, 91: 1617-1627.

Rousset F (2000). Genetic differentiation between individuals. Journal of Evolutionary Biology 13:58-62.

Sedda L, Brown HE, Purse BV, Burgin L, Gloster J, Rogers DJ (2012). A new algorithm quantifies the roles of wind and midge flight activity in the bluetongue epizootic in northwest Europe. Proceedings of the Royal Society B-Biological Sciences, 279: 2354-2362.

Sedda L, Morley D, Brown HE (2015). Characteristics of Wind-Infective Farms of the 2006 Bluetongue Serotype 8 Epidemic in Northern Europe. EcoHealth, 12: 461-467.

Shults P, Ho A, Martin EM, McGregor BL, Vargo EL (2020). Genetic Diversity of *Culicoides stellifer* (Diptera: Ceratopogonidae) in the Southeastern United States Compared With Sequences From Ontario, Canada. Journal of medical entomology, 57: 1324-1327.

Tajima F (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics, 123: 585-595.

Wegmann D, Leuenberger C, Neuenschwander S, Excoffier L (2010). ABCtoolbox: a versatile toolkit for approximate Bayesian computations. Bmc Bioinformatics, 11.