

Histoire évolutive de *Parnassius Apollo* par des inférences démographiques à l'échelle des principaux massifs Français



Résumé

Mieux connaître l'histoire d'une espèce et ses réponses démographiques aux changements climatiques passés, peut aider à sa conservation dans un contexte de réchauffement climatique. Cette étude s'intéresse dans un premier temps à la structure génétique d'un papillon emblématique des montagnes en Europe, *Parnassius Apollo*, puis dans un second temps à inférer son histoire évolutive.

L'échantillonnage est constitué de 268 individus provenant des principaux massifs français (83 localités), génotypés pour 3899 loci (ddRADseq). La population est structurée en 6 groupes génétiques correspondant aux principaux massifs : Auvergne, Ardèche/Cévennes, Pyrénées, Jura/Alpes du Nord, Chartreuse et Alpes du Sud. Les analyses montrent la présence de barrières migratoires notamment dans la vallée du Grésivaudan dans les Alpes ou dans le massif central conduisant à un isolement de certaines populations (le plus marquant étant l'Auvergne). Ensuite, avec les inférences démographiques, il apparaît clairement que ces 6 groupes génétiques ont une histoire similaire, expansion durant le refroidissement entre le LIG et le LGM (papillon adapté au froid), puis déclin au moment du réchauffement entre le LGM et le mid-holocène. A l'intérieur de chaque massif, le déclin est plus ou moins graduel et prononcé, avec un déclin marqué dans le massif Central (Auvergne, Ardèche et Cévennes), plus modéré dans les Pyrénées, et une diversité génétique et une connectivité qui restent importantes au sein des massifs alpins.

Abstract

Knowing more about the history of a species and its demographic responses to past climate change can help in its conservation in a context of global warming. This study first focuses on the genetic structure of an emblematic European mountain butterfly, *Parnassius Apollo*, and then on inferring its evolutionary history. The sample consists of 268 individuals from the main French massifs (83 localities), genotyped for 3899 loci (ddRADseq). The population is structured in 6 genetic groups corresponding to the main massifs: Auvergne, Ardèche/Cévennes, Pyrenees, Jura/Northern Alps, Chartreuse and Southern Alps. The analyses show the presence of migratory barriers, notably in the Grésivaudan valley in the Alps or in the Massif Central, leading to the isolation of certain populations (the most striking being the Auvergne). Then, with the demographic inferences, it appears clearly that these 6 genetic groups have a similar history, expansion during the cooling between the LIG and the LGM (cold-adapted butterfly), then decline at the time of the warming between the LGM and the mid-Holocene. Within each massif, the decline is more or less gradual and pronounced, with a marked decline in the Massif Central (Auvergne, Ardèche and Cévennes), a more moderate decline in the Pyrenees, and a genetic diversity and connectivity that remains important within the Alpine massifs.

Sommaire

❖ Introduction

❖ Matériel & méthode

- Présentation des données
- Génotypage des SNPS
- Analyse de diversité génétique et structure de population
- Inférences démographiques

❖ Résultats

- Structure génétique
- Isolation par la distance
- Inférences démographiques

❖ Discussion

- Une population structurée par massif
- Une histoire évolutive commune
- Limites et perspectives

❖ Conclusion

Introduction

Mieux connaître l'histoire d'une espèce et ses réponses démographiques aux changements climatiques passés peut aider à sa conservation dans un contexte de réchauffement climatique. Ces indices sur le passé constituent un socle de connaissances fondamentales permettant la compréhension de la situation actuelle ainsi qu'être une base pour la prédiction d'événements futurs (Todisco et al., 2010). Cette histoire évolutive peut souvent être résumée par des changements de taille de populations, qui en fonction de la distribution géographique, peuvent conduire à des populations plus ou moins isolées. Ainsi en prenant en compte une notion de migration inter-population, il est alors possible de définir le rôle relatif de la dérive génétique et de la sélection dans le façonnement de la différenciation génétique intra et inter-population et de la structure de la population (Ometto et al., 2015). Il a notamment été reconnu que les variations climatiques sont un des principaux facteurs influençant les schémas évolutif et biogéographique d'espèces dans le monde entier (Hewitt, 2004). On peut distinguer alors une alternance entre des périodes de refroidissement et de réchauffement conduisant à une expansion ou une contraction des espèces en fonction de leur besoin écologique (Teixeira et al., 2021). Ces oscillations façonnent alors l'histoire démographique d'une espèce. La plupart des études s'intéressant aux inférences démographiques, se concentre sur des fluctuations climatiques bien prononcées comme notamment le dernier interglaciaire (LIG, il y a environ 112-132 000 ans), le dernier maximum glaciaire (LGM, il y a environ 19-26 500 ans) ainsi que le mid-holocène, période ayant un fort réchauffement (~6000 ans) (Kebaïli et al., 2022; Kukla et al., 2002).

L'impact de ces oscillations est d'autant plus vrai pour des insectes, hautement sensibles aux changements environnementaux en ayant des besoins écologiques très spécialisés, comme notamment les papillons, un des groupes d'organismes dont la répartition a été la plus étudiée à travers le temps (Todisco et al., 2010).

Cependant en ce qui concerne les espèces alpines, une grande part de mystère est présente et plusieurs scénarios évolutifs se confrontent (Kebaïli et al., 2022). En effet, il a été admis que pendant la glaciation, de nombreux taxons étaient limités à des refuges situés dans le sud et libérés de la glace (Hewitt, 1996; Taberlet et al., 1998) et que la distribution actuelle des espèces européennes est le résultat d'une colonisation de ces refuges vers le nord (Strandberg et al., 2011). Ce scénario aurait pour principale conséquence un fort goulot d'étranglement pendant les périodes glaciaires or ceci n'est pas forcément vérifié pour les espèces alpines qui peuvent avoir survécu localement dans des refuges extra-méditerranéens (Schmitt, 2007, p.

2). Ces même périodes de refroidissement pourraient même causer l'expansion d'espèces adaptées au froid ainsi qu'améliorer la connectivité entre les populations de différentes chaînes de montagnes qui sont mises en contact dans les zones libres de glace à basse altitude contrebalançant ainsi les effets de la dérive (Kebaïli et al., 2022). Ce scénario nommé modèle de l'archipel alpin refuge, suppose que la dispersion entre les différentes populations a lieu au moment d'une expansion. Ce scénario est contredit par un autre, celui du modèle de radiation et de fragmentation ancestrale expliquant qu'une population ancestrale à coloniser simultanément les habitats montagnards pendant une période interglaciaire suivie d'un isolement sur les différents massifs après la contraction (Schoville et al., 2011).

C'est pour cela que cette étude a pour but de déterminer l'histoire évolutive d'une espèce de papillons emblématique des montagnes et adaptée au froid, *Parnassius apollo*, à l'échelle des principaux massifs français. Avec ici comme principal objectif de pouvoir éclaircir les différents mystères autour de son apparition et sa répartition en France. Par exemple, à l'aide de données mitochondriales, une théorie a été évoquée expliquant que la répartition actuelle de *P. apollo* est la résultante d'une colonisation vers l'ouest depuis l'Asie (Todisco et al., 2010). Les questions qui restent donc en suspens sont alors : comment sont génétiquement structurées les différentes populations de *P. apollo* française ? Quels scénarios collent le mieux pour son histoire évolutive (modèle des refuges de l'archipel alpin vs radiation et fragmentation ancestrale) ? Est-ce que le scénario d'une colonisation est retrouvé avec des données ddRAD ? Cette étude découle notamment d'une étude précédente réalisée par Kebaïli & al, ayant des objectifs très similaires, mais plus concentrée sur le massif alpin et quelques individus dans le puy de Sancy. Ici en plus d'avoir plus d'individus dans les alpes et l'auvergne, d'autres massif sont pris en compte. En effet, cette étude s'étend sur 83 localités dans différents massifs comme Pyrénées, l'Auvergne, l'Ardèche, les Cévennes, les Alpes comprenant les Alpes du sud, la Chartreuse, Vercors, Belledonne, les Bauges ainsi que le Jura pour un total 283 individus. L'ensemble de ces individus est le résultat de deux périodes d'échantillonnage, une en 2018 et une en 2021 et a été séquencé à partir d'un prélèvement d'une patte à l'aide d'un protocole ddRAD. Pour répondre aux enjeux de cette étude, cette analyse va être séparée en deux grandes parties, une première partie centrée sur la structure génétique actuelle et la corrélation entre les données génétiques et les données géographiques relevées. Ensuite une seconde partie sur les inférences démographiques avec 3 méthodes complémentaires, le stairway plot permettant l'estimation d'histoire évolutive de population sur la base de spectre de fréquences en faisant une assumption sur le taux de mutation (Liu and Fu, 2015). La seconde sera une méthode ABC (DIYABC) permettant d'éviter

l'assomption d'un taux de mutation pouvant être source d'imprécision , et de comparer plusieurs scénarios plus complexes. Cependant un grand problème de ces méthodes est la non-prise en compte de taux migration ayant pour conséquence des scénarios biaisés (Cornuet et al., 2014, 2010). Pour cela une dernière méthode d'inférence démographique sera utilisée, nommée DILS (Demographic inference with linkage sélection) basé aussi par une méthode ABC mais ici en comparant des modèles comprenant des taux de migrations (Fraïsse et al., 2020).

Matériel & méthode

Présentation des données

Les données utilisées dans cette étude proviennent de deux campagnes d'échantillonnage (Figure.sup 1), une première lancée en 2018 sur 107 individus principalement basés sur le massif alpin et quelques individus provenant du puy de Sancy. La seconde campagne lancée en 2021 sur 176 individus qui en plus de mieux couvrir les massifs alpins et auvergnat, ajoute aussi de nouveaux massifs comme les Pyrénées, l'Ardèche et les Cévennes. Au total, le jeu de données est constitué de 283 individus sur 83 localités différentes qui ont été séquencé par DDrad, en utilisant Sbf1 et Msp1 comme enzyme de restriction.

Génotypage des SNPs

Avec le séquençage ddrad, environ 734 000 000 de reads on été obtenus, dans lequel les séquences correspondant aux adaptateurs on été éliminés à l'aide de l'outil bbdutk provenant du package BBmap (Bushnell, 2014) et l'appel des snps a été réalisé avec le pipeline STACKs basé sur un mapping sur génome de référence. Les différents reads ont été démultiplés, filtrés et tronqués à 110 bp à l'aide de la fonction process_radtag. 11 individus ayant moins de 100 000 reads ont été exclus de l'analyse (Tableau supplémentaire 1). Aussi Msp1 ayant un site de restriction plutôt court (4 pb) et étant principalement constitué de GC, il s'avère difficile de retrouver les reads provenant de son site (R2) causant alors une forte quantité de R1 n'ayant pas son read complémentaire (appelé Rem1). Afin de limiter cette perte importante de données, une concaténation des R1 ayant trouvé un R2 et des Rem1 a été réalisée. Cette analyse étant basée sur deux campagnes d'échantillonnage, 4 individus ont pu être rééchantillonnés, nécessitant une étape de rassemblement des reads pour chaque correspondance. Ici l'étude est basée sur un génome référence, publié par Podsiadlowski & al, semblant plutôt satisfaisant avec notamment 97% des gènes BUSCO de lépidoptères

retrouvés complets en copie unique (Podsiadlowski et al., 2021) . La fonction `bwa index` a permis d'indexer ce génome de référence suivie de la fonction `bwa mem` ne gardant que les reads ayant une qualité de mapping supérieur ou égale à 20. Enfin un catalogue des loci pour tous les individus a été créé à l'aide de la fonction `gstacks` permettant la production du jeu de données à l'aide de la fonction `populations`. À savoir que seulement les snps étant présents dans au moins 60% des individus échantillonnés, ayant une couverture moyenne supérieure à 10. Les allèles étant présents qu'une à deux fois dans tout le jeu de données sont enlevés, ainsi que les locus et les individus ayant plus de 40% de données manquantes. Un snp par locus a été conservé, conduisant donc à un dataset de 3899 loci indépendants pour 268 individus. En fonction des différentes analyses réalisées, d'autres filtrations additionnelles ont pu avoir lieu.

Analyse de diversité génétique et structure de population

Concernant les analyses de diversité génétique et la structure de population, un filtre additionnel a été réalisé en ne sélectionnant que les snp ayant une maf (fréquence de l'allèle mineur) supérieur à 1%. Pour chaque localités ayant au minimum 2 individus (Nazareno et al., 2017), différentes mesures ont été évaluées comme le nombre d'hétérozygotes observés (H_o), d'hétérozygotes attendus (H_e), la déviation par rapport à l'hétérozygotie attendus due à un accouplement non aléatoire (F_{is}) ainsi que la différenciation de chaque localité par rapport à l'ensemble (F_{ST}) ou des F_{ST} inter-population (Pairwise F_{ST}) à l'aide du package Hierfstat (Goudet, 2005). Les différents intervalles de confiance ont été mesurés pour chacune de ces mesures à l'aide de 1000 bootstraps à travers les loci et 100 pour les F_{ST} par paires. Avec ces mesures de diversité génétique il est possible d'estimer le D_{Tajima} , une mesure permettant de savoir si l'évolution d'une population est due au hasard (hypothèse neutre), soit par un processus non aléatoire tel qu'une contraction, valeur positive, ou une expansion, valeur négative (Tajima, 1989). Le principe de cette mesure est d'estimer la différence entre le nombre de différence nucléotidiques estimé à partir d'une comparaison par paire ainsi que le nombre de sites qui se ségrégent. Sous l'hypothèse nulle, il est attendu que cette différence soit nulle avec une population de taille constante. Dans cette étude, cette mesure est estimée à partir des haplotypes à l'aide `vcftools -dtajima` avec une fenêtre glissante de 100kb.

La structure génétique de la population a été analysée à l'aide d'une analyse en composante principale (ACP) en utilisant le package PCadapt. Ensuite le K correspondant au nombre de clusters génétiques distinct a été inféré à l'aide de 3 méthodes différentes, SNMF, DAPC et

Admixture, en choisissant pour chaque, la valeur de cross-entropie la plus faible pour un k allant de 1 à 12 (Figure.sup 2). Une carte liant les résultats d'admixture avec les données géographiques indiquées par les collecteurs a été réalisée à l'aide du logiciel Qgis.

De plus, une autre méthode nommée FineRAD structure (<https://github.com/millanek/fineRADstructure>) a été utilisée permettant aussi d'estimer une structure de population mais cette fois-ci à l'aide des haplotypes (Malinsky et al., 2018). Cette méthode est notamment capable d'estimer des matrices de coancestries avec des données de séquençage RAD couplé avec un algorithme de clustering basé sur des chaînes de MCMC. Pour fonctionner elle nécessite un fichier d'entrée sous format radpainter obtenus grâce à la fonction population de stacks et la figure est produite via R.

La relation entre la distance génétique et la distance géographique a aussi été testé dans chaque massif à l'aide de pairwise FST intra-population estimé à l'aide du package hierfstat. A noter que pour toutes les analyses utilisant des coordonnées géographiques le jeu de données utilisé est constitué du dataset de départ moins 10 individus des Pyrénées qui ont des informations géographiques inconnues (Problème d'étiquetage au moment de l'envoi des échantillons). Pour ces analyses, le jeu de données est constitué alors de 3899 locus pour 258 individus. Les distances géographiques ont été estimées à l'aide de la fonction distm provenant du package geosphere (Hijmans, 2021). Le Fst a été normalisé par la formule: $Fst/(1 - FST)$ permettant la réalisation de test de Mantel avec 10 000 000 de permutation grâce à la fonction mantel.randtest du package ade4 (Dray, 2007). Ce test de mantel n'as pas été appliqué aux groupes de l'Ardèche et des Cévennes ayant une trop faible quantité de données (2 et 1 sites échantillonnés respectivement)

Une autre mesure d'isolation par distance a été réalisé à l'aide du package EEMS (<https://github.com/dipetkov/eems>) permettant de comparer l'isolation ainsi que la migration observée entre chaque massif par rapport à ce que l'on attend avec le hasard (Petkova et al., 2016). Pour appliquer EEMS, il est nécessaire de créer un objet géographique, shapefile de la zone d'intérêt (ici grand sud de la France) via QGIS, les données géographiques des localités et enfin les données génétiques associées à ces populations. Cette méthode utilise une MCMC pour l'estimation de ces paramètres, ici dans cette étude, 3 différentes chaînes sont faites avec une phase de burnin de 1500000 itérations. Aussi pour estimer la valeur du taux de migration et de la diversité génétique par zone, cette méthode se base sur l'établissement d'une grille avec un maillage paramétrable, ici correspondant à 1000 demes. Les résultats seront interprétés si et seulement si les 3 différentes chaînes convergent ensemble.

Inférence démographique

L'histoire évolutive a été déterminée par 3 méthodes d'inférence démographique différentes. Pour ces analyses, le jeu de données est constitué du jeu de données global mais sans maf pour ne pas éliminer trop de singleton, témoin de l'histoire récente. De plus, pour chaque individu, une attribution à un cluster génétique est attribuée (6 groupes génétiques correspondant aux différents massifs déterminés par les analyses de structure). Pour comparer l'histoire évolutive de deux populations différentes, il est préférable d'avoir des jeux de données de taille similaire. Or ici les tailles de chaque population sont assez semblables entre 13 et 40 sauf pour le groupe des Alpes du Sud ayant 92 individus. Pour y remédier, le groupe Alpe du sud a été subdivisé en 3 groupes : massif de Vercors et de Belledonne (Ver_Bell), les populations du nord des Ecrins (As_N) et les populations du sud de la région PACA (AS_S). Premièrement la démographie a été estimée pour ces populations séparées à l'aide du Stairway plot (Liu and Fu, 2015), qui permet l'estimation de changement continu de la taille efficace d'une population au cours du temps en se basant sur un spectre de fréquence de SNPS (SFS). Dans le cas de l'étude, aucune information n'était disponible sur le caractère ancestral ou dérivé des allèles, obligeant donc l'emploi d'un SFS dit replié (Booker et al., 2017). Aussi cette méthode nécessite l'assomption d'un taux de mutation, or de précédentes études sur ce papillon ont fixé un taux de mutation de $2e-9$ (Kebaïli et al., 2022; Keightley et al., 2015) en sachant que pour celui-ci, une génération est équivalente à une année. De par la nature de cette méthode, la présence de données manquantes est incompatible, elles sont alors approchées à l'aide de la fonction impute de LEA en prenant en compte 6 groupes génétiques distincts.

Dans un second temps, le logiciel DIYABC (approche coalescente utilisant ABC) a été utilisé permettant de comparer différents scénarios démographiques sans faire d'assomption sur le taux de mutation (Cornuet et al., 2014). De plus les données manquantes ne sont pas limitantes ici permettant de prendre le jeu de données sans imputation. À l'aide de cette méthode, plusieurs scénarios ont été confrontés. En suivant l'approche réalisée avec le stairway plot, des premiers scénarios en population isolée ont été réalisés. Pour chaque population 6 scénarios démographiques ont été comparés: (1) taille constante, (2) expansion entre le LIG et LGM puis déclin entre le LGM et mid-holocène (déclin ancien), (3) expansion entre le LIG /LGM puis déclin entre le mid-holocène et le présent (déclin récent), (4) expansion entre le LIG /LGM puis deux déclins (déclin ancien et récent), (5) expansion entre le LIG /LGM et enfin (6) un déclin entre le LGM et mid-holocène. Pour chaque modélisation avec ABC, des gammes de priors ont besoin définies. En mono population, les priors de

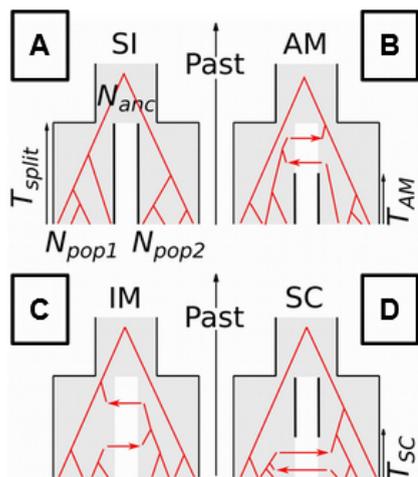
temps correspondent aux SDM réalisés dans l'étude de Kebaili & al et les priors de taille proviennent des résultats du stairway plot. 100 000 simulations par scénario ont été réalisées en utilisant comme statistique résumé la moyenne de la diversité génétique dans chaque population. Les probabilités postérieures de chaque scénario, sont calculées à l'aide d'une régression logistique des 1% de données simulés les plus proches des données observés. La confiance dans le choix de scénario a été estimée à l'aide des erreurs de type 1 et de type 2 grâce à 100 nouvelles données simulées à partir des priors, calcul réalisé sur R grâce au package abcrf.

Dans un second temps, des scénarios de divergence entre les différents massifs ont été comparés. Pour cela, des scénarios comparant les populations deux à deux sont réalisés. En sachant qu'avec 6 clusters génétiques, cela correspondrait à 15 couples différents ce qui semble assez conséquent. Or dans les 6 clusters génétiques, 3 proviennent des alpes avec une histoire en mono population similaire (Stairway plot & DIYABC). Pour limiter les couples, le groupe Ver_Bell (Alpes du sud) et B_HS_J (Alpes du nord) seront les représentants conduisant maintenant à 10 couples à tester. Un dernier couple sera étudié comprenant l'Ardèche et les Cévennes. Pour chaque couple, deux scénarios ont été comparés, avec un scénario ayant un split des deux populations pendant l'expansion entre le LIG et le LGM, ainsi qu'un deuxième scénario ayant toujours une expansion entre le LIG et le LGM mais ayant un split pendant le déclin entre LGM et le présent. Les probabilités estimées par DIYABC permettant le choix de chaque scénario ont été rassemblé pour chaque test (Table sup 2)

Pour les scénarios ayant plusieurs populations, de nouvelles statistiques résumées s'ajoutent à la diversité génétique comme des distances de FST et NEI permettant de réaliser le choix de modèle.

Enfin la dernière méthode employée est nommée DILS permettant des inférences démographiques en prenant en compte des taux de migration. Pour cette méthode seulement des comparaisons entre deux populations sont possibles. 4 scénarios sont testés à l'aide d'un algorithme ABC : Stricte isolation, isolation avec une migration ancienne, isolation avec migration où enfin un contacte secondaire (isolation stricte puis migration récente).

Schémas 1 : Scénario testé avec DILS



Schémas des différents scénarios comparés par la méthode DILS, avec comme scénario (A) une isolation stricte, (B) une isolation avec une migration ancienne, (C) une isolation avec migration continue et (D) une isolation avec un contact secondaire récent.

Légende : N_{pop} : taille de population récente, N_{anc} : Taille de la population ancestrale, T_{split} : Temps de split, T_{AM} : Temps ou a eu lieu la migration ancienne, T_{SC} : Temps ou a eu lieu le contacte secondaire

Cette méthode est employée avec les mêmes couples de populations que DIYABC, mais avec cette fois-ci des séquences fasta par locus (Sortie –fasta sample de population). Utilisant elle aussi un modèle ABC, cette technique nécessite différents priors, pour la taille de pop [1000-1000000], pour le temps de split [0-150 000] ans, pour la migration [0.4%-20%] et un taux de

mutation et de recombinaison intralocus. Sachant que pour toutes nos inférences démographiques, le jeu de données a été filtré en prenant seulement 1 snp par locus ce qui est impossible ici à faire et que l'on a en moyenne 6.5 snp par locus dans le jeu de données d'origine, nous avons choisi un taux de mutation à $1e-8$ ($2e-9$ au départ) pour simuler cette filtration. Pour appliquer cette méthode, il est nécessaire d'avoir des groupes de taille égale, ce qui est réalisé avec un sous-échantillonnage de 13 individus dans chaque population.

Résultats

Concernant les différentes mesures de diversités génétiques estimées pour les 83 différentes localités, il apparaît que pour l'hétérozygotie observée et attendue atteint des valeurs plus importantes pour les massif alpins que pour les autres massifs avec en moyenne une valeur de 0.094 et de 0.13 contre 0.066 et 0.087 pour le massif central ainsi que 0.083 et 0.121 pour les Pyrénées (Table sup 1). Pour le H_e , les différences observées entre chaque massif sont significatives ($p < 0.004$ avec un test de kruskall wallis). Le FIS semble similaire entre les Alpes, le massif central et les Pyrénées. Cependant un fait marquant que le FIS est significativement supérieur à 0 dans toutes les localités soulignant qu'aucune de ces populations est en équilibre démographique et donc la présence de consanguinité. Pour la différenciation génétique estimée à l'aide du F_{ST} spécifique de chaque localité, le massif central est alors assez remarquable avec une valeur importante de 0.42 avec notamment l'Auvergne dépassant 0.5 comparés aux massifs Alpains et aux Pyrénées ayant une valeur similaire, 0.18 et 0.24 respectivement (différence significative $p < 0.02$, Table sup 1). Au niveau du D_{Tajima} , seules les populations Alpines semblent avoir une valeur négative (valeur moyenne dans les Alpes de -0,492 contre 0.44 dans le Massif Centrale et 0.24 dans les

Pyrénées). Néanmoins, les différences entre chaque massif ne sont pas significatives, et ne permettent pas de rejeter l'hypothèse nulle avec des intervalles de confiance se chevauchant et comprenant 0.

Structure génétique:

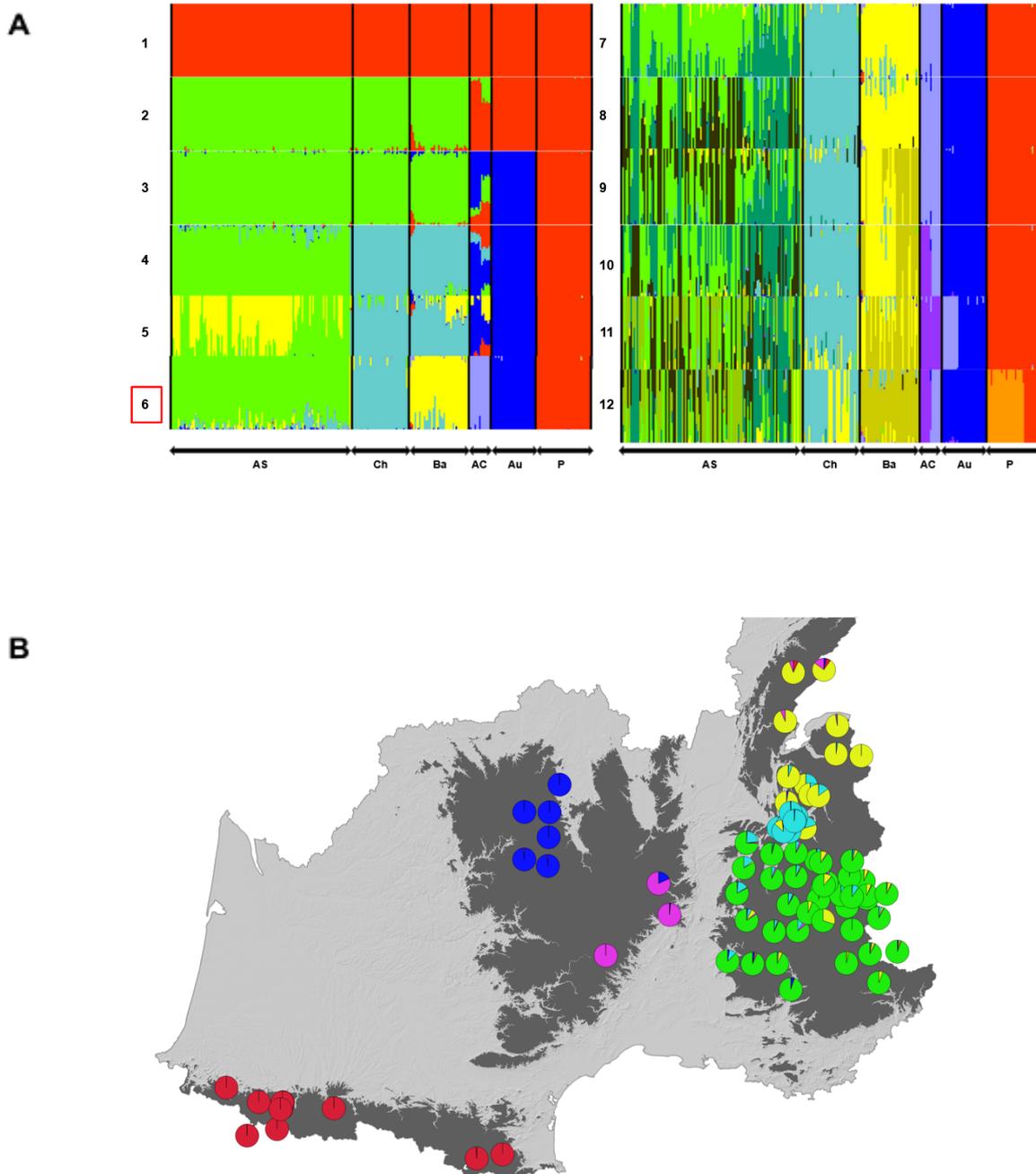


Figure 1 : Structure génétique des populations de P.apollo dans le massif des Alpes, l’Auvergne, l’Ardèche, les Cévennes et les Pyrénées

(A) Structure génétique pour un k allant de 1 à 12 pour tous les individus réalisée avec admixture ainsi que (B) une carte représentant le taux d’assignation moyen par localité pour un $K = 6$ réalisé sur QGIS.

Légende : AS : Alpes du sud , Ch : Chartreuse, Ba : Bauges_Haute-Savoie_Juras, AC: Ardèche-Cévenne, AU : Auvergne, P : Pyrénées

Le K optimale pour ces 268 individus (Fig.sup 2) varie en fonction de la méthode. En effet pour DAPC la valeur est de 3 clusters, tandis que pour SNMF la valeur serait de 5 alors que pour Admixture le K optimale pourrait être 3,4 ou 5 (Valeur minimale égale pour ces 3 valeurs). En regardant la structuration de la population avec un K allant de 1 à 12 avec admixture (Fig1A), un schéma semble se dessiner. Premièrement, il y a une distinction très nette entre les alpes et le reste des massifs à K = 2 avec le groupe comprenant l'Ardèche et les Cévennes étant très métissé, devenant un groupe bien distinct à partir de K = 6. A ce K (Fig1A), une forte structuration par massif est remarquable avec un groupe Pyrénées, un groupe pour l'Auvergne, un groupe avec Ardèche plus Cévennes, ainsi que les Alpes séparés en 3 différentes parties, les alpes du sud, le massif de la Chartreuse et les alpes du nord comprenant le massif des Bauges, la Haute-savoie et le Jura. A noter aussi que ces groupes semblent être plutôt conservés avec l'augmentation du K sauf pour les alpes du sud ayant un départ d'individualisation à partir du k = 7.

Pour compléter et vérifier cette structure, les haplotypes de ces individus ont été analysés à l'aide la méthode de FineRad Structure (Figure sup 3). Il est alors facilement distinguable les 6 groupes trouvés précédemment avec les individus les plus similaires représentés avec une couleur plus sombre. La séparation entre les alpes semble beaucoup moins nette que pour les séparations entre les Pyrénées l'Auvergne et Ardèche/Cévenne . De plus le groupe Ardèche/Cévenne avec les haplotypes semble assez mélangé avec le groupe de l'auvergne.

Ensuite, en gardant cette idée, il est maintenant possible de comparer la différenciation génétique entre chacun de ces grands groupes. Pour cela, une comparaison de FST deux à deux est réalisée (Table 1).

	Pyrénées	Auvergne	AC	AS	BHSJ	Chartreuse
Pyrénées	NA	0.31 [0.28-0.32]	0.19 [0.17-0.20]	0.19 [0.17-0.20]	0.17 [0.16-0.18]	0.18 [0.17-0.19]
Auvergne	0.31 [0.28-0.32]	NA	0.24 [0.22-0.26]	0.25 [0.23-0.26]	0.25 [0.24-0.27]	0.26 [0.24-0.28]
AC	0.19 [0.17-0.20]	0.24 [0.22-0.26]	NA	0.16 [0.14-0.17]	0.15 [0.13-0.16]	0.16 [0.14-0.17]
AS	0.19 [0.17-0.20]	0.25 [0.23-0.26]	0.16 [0.14-0.17]	NA	0.03 [0.02-0.04]	0.03 [0.02-0.04]
BHSJ	0.17 [0.16-0.18]	0.25 [0.24-0.27]	0.15 [0.13-0.16]	0.03 [0.02-0.04]	NA	0.02 [0.01-0.03]
Chartreuse	0.18 [0.17-0.19]	0.26 [0.24-0.28]	0.16 [0.14-0.17]	0.03 [0.02-0.04]	0.02 [0.01-0.03]	NA

Table 1 : Différenciation génétiques à l'aide de pairwise fst

Il est alors notable que la différence maximale est atteint entre l'Auvergne et les Pyrénées

atteignant 0.31. Globalement, l'Auvergne a pour chaque couple, une valeur de FST couplée supérieure à 0.20, étant significativement supérieures à tous les autres couples (estimé à l'aide des intervalles de confiance). Les FST minimales sont obtenus entre les groupes provenant des Alpes, avec une valeur variant entre 0.02 et 0.03. Le groupe AC semble être à la même distance génétique avec les Pyrénées qu'avec les groupes alpins (FST variant de 0.15 à 0.19 avec une différence non significative).

Isolation par la distance:

Pour finaliser la structure génétique des populations de *Parnassius apollo* française, une analyse d'isolation par la distance a été faite, en recherchant si la corrélation entre la distance géographique et la distance génétique est plus ou moins forte qu'attendus si il y a seulement présence d'isolation par la distance (Figure 2). En commençant par l'isolation par la distance, une grande zone ayant un indice négatif entre les massifs alpins et l'Ardèche plus les Cévennes ainsi qu'entre les alpes du nord et du sud avec une séparation dans la vallée du Grésivaudan. De plus l'auvergne semble être entouré de zone ayant un indice similaire contrasté par une indice intra-Auvergne neutre. Au sein des alpes du nord, des alpes du sud et des Pyrénées, il semble avoir un indice assez positif (Figure 2a). La méthode EEMS permet aussi d'estimer des indices de diversité génétiques. Comme à l'instar de nos précédentes mesures, les massifs alpins semblent être la zone ayant un indice de diversité génétique le plus fort suivie des Pyrénées et complétés par le massif central qui, pour ceux-ci, présente un indice fortement négatif (Figure 2B).

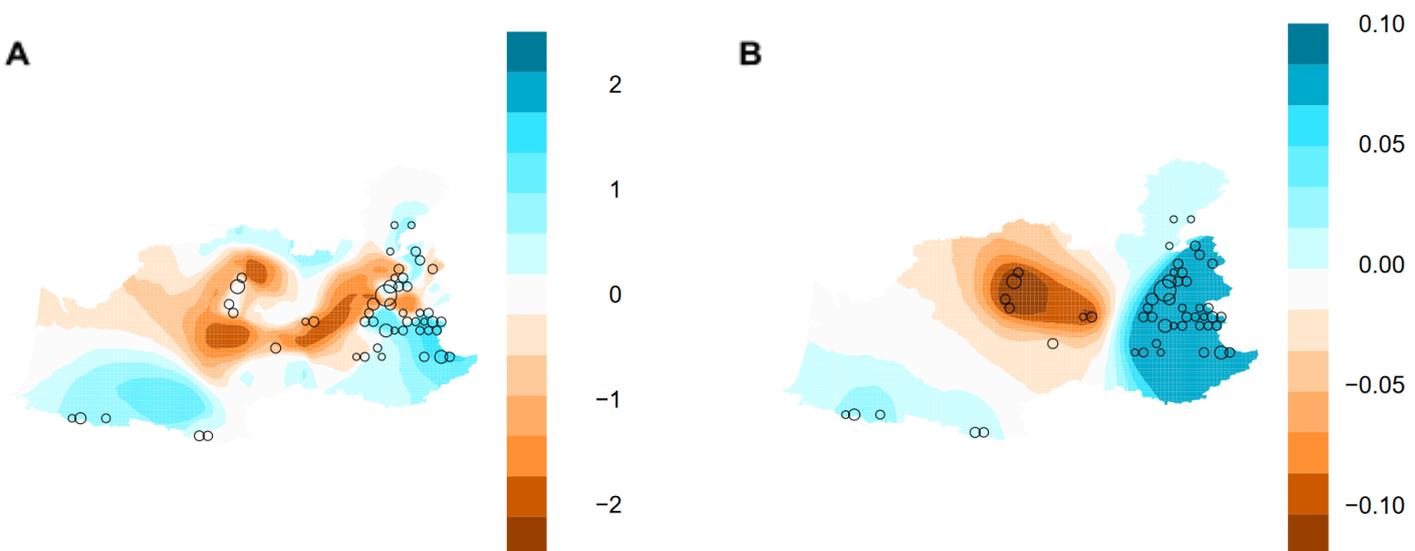


Figure 2: Isolation par la distance

(A) Surface effective de migration permettant de visualiser à l'échelle de la zone d'étude les taux de migration supérieurs (en bleu) ou

inférieurs (en rouge) à l'attendu sous l'hypothèse d'isolement par la distance (B) estimation de la distribution spatiale des taux de diversité génétique obtenus à l'aide de EEMS avec une grille de 1000 demes.

Des tests de Mantel ont été réalisés par localités pour vérifier la significativité de la présence ou non d'isolation par la distance ainsi que d'estimer la part explicative de celle-ci sur nos données (Figure sup 3) . Pour les sites alpins, Auvergnat et Pyrénéen, le test de mantel est significatif (p valeur égale à $9e-8$, 0.003 et $1e-4$ respectivement) cependant la part de la variation génétique expliquée par l'isolation par la distance varie. En effet pour les Alpes, le R estimé est au alentour de 0.24 (Figure sup 3) ce qui est plutôt faible comparé aux Pyrénées, égale à 0.79 (Figure sup 3) et l'auvergne étant la valeur maximale avec $R = 0.96$ Figure sup 3).

Inférence démographique

En seconde partie de cette analyse, des inférences démographiques ont été réalisées. Les inférences ont été réalisées par 3 méthodes différentes:

La première est l'application du stairway plot (Figure 3) pour les groupes génétiques identifiés précédemment avec ici un subset dans les alpes du sud pour lisser le nombre d'individus entre les différentes populations, fixé à 13 individus, taille de la population la plus petite, le groupe Ardèche/Cévennes. Il en ressort ici des valeurs plutôt similaires pour chacune des courbes avec une taille ancestrale datant de plus 130 000 ans comprise entre 25 000 et 50 000 individus. Ensuite les courbes entre dans une phase de croissance entre 100 000 et 120 000 ans avant maintenant, étant un peu plus récent pour deux sous-échantillons des alpes du sud avec AS_S entrant en phase croissante aux alentours de 50 000 ans avant maintenant et AS_N vers 75 000 ans. Après cette croissance une phase de décroissance lente à eu lieu pour toutes les populations jusqu'à environ moins de 2000 ans avant maintenant pour ensuite entrer dans une phase de décroissance forte. À noter que le groupe de l'Auvergne et de l'AC sont les deux premières populations à entrer dans cette phase de déclin (~10 000 ans avant maintenant pour AC et ~5000-6000 ans pour l'Auvergne).

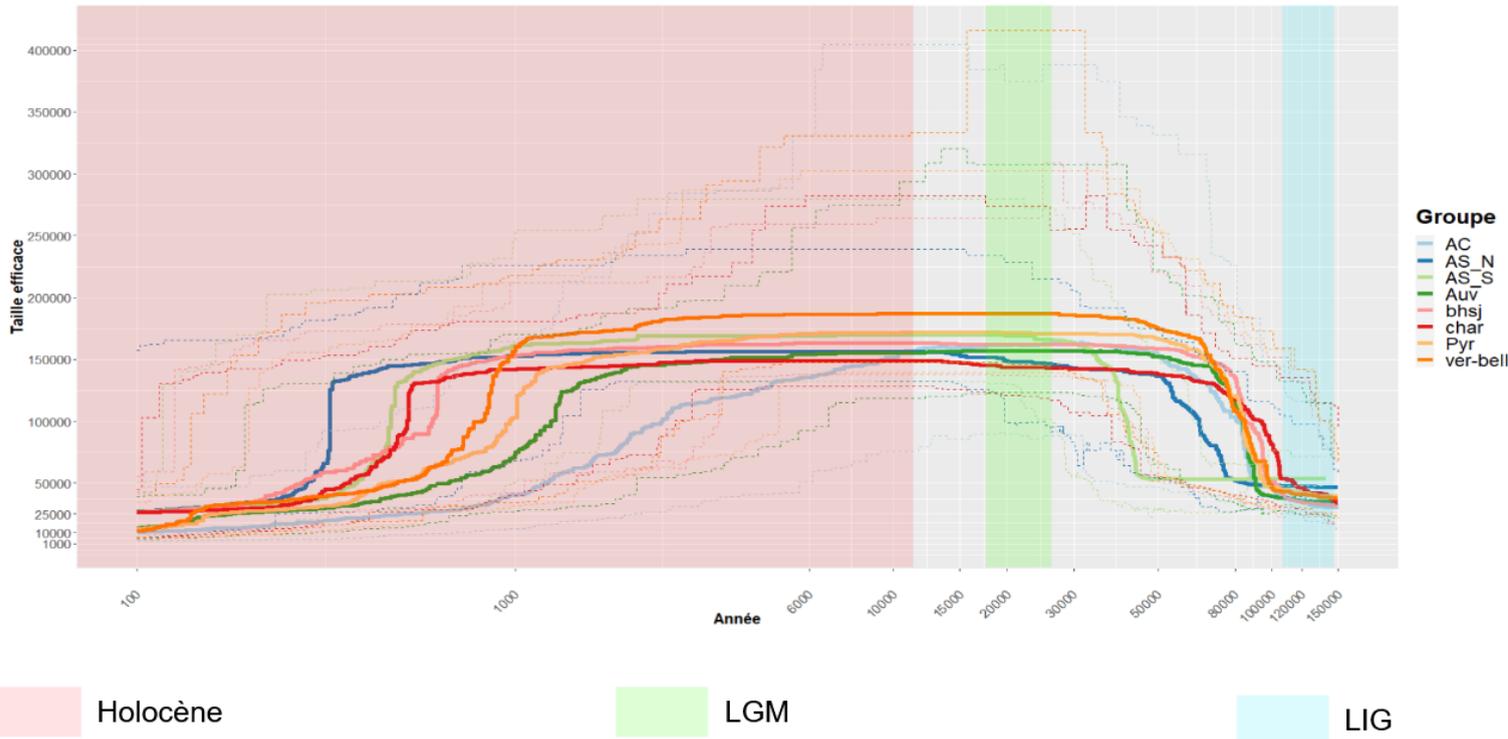


Figure 3: Changement continu de la taille efficace de chaque population

Cette décroissance forte a eu lieu entre 1500 et 300 ans environ et fait varier la taille efficace de chaque population d'environ 150 000 individus à moins de 50 000. Dans cette figure deux grandes phases de divergence entre les différentes populations, une plutôt faible pendant l'expansion avec une histoire qui semble différente pour deux sous-groupes des alpes du sud, et une divergence plutôt forte pendant la décroissance avec ici une sorte d'éclatement. Il est à préciser ici, qu'en prenant en compte les intervalles de confiance de chacune de ces histoires, il semble alors très compliqué de distinguer des différences significatives.

Cette histoire évolutive est globalement concordante avec l'approche en mono-population avec DIYABC pour les 8 groupes une histoire similaire avec une phase d'expansion il y a environ 80-90 000 ans avant maintenant et une phase de déclin brutal depuis 12 000 ans (Table sup 2, Figure sup 4). Dans cette approche, il est à noter que les modèles de tailles constantes ont pour toutes les populations était fortement rejetée avec une valeur de post-probabilité de 0 pour tout (Table sup 2). De plus pour aussi tester les scénarios décrits dans le papier de C.Kebaili & al, des modèles faisant la nuance entre un déclin dit "ancien" entre 6000 et 20 000 ans et un déclin dit "récent" avant 6000 ans avant maintenant ont été réalisés. Des sortes de patterns sont alors identifiables par massif, notamment avec les

Pyrénées et les Alpes du nord avec une expansion suivie de deux déclin, les populations du massif central ayant une expansion suivie d'un déclin ancien et enfin les alpes du sud ayant une expansion suivie d'un déclin récent. Une exception est présente pour AS_N semblant suivre le pattern des populations des alpes du nord (Figure sup 4). Cependant avec ces modèles, il est alors remarquable qu'entre un scénario avec soit un déclin récent, soit un déclin ancien soit encore les deux déclin, la distinction et le choix de modèle semblent assez complexes conduisant donc à de fort taux d'erreurs. Globalement, les erreurs de type 1 sont assez conséquentes, allant de 11% jusqu'à 65% et avec une moyenne de 53% d'erreur. Concernant les erreurs de type 2, en regardant la moyenne entre les scénarios faisant intervenir un déclin ou plus, il apparaît alors une probabilité de se tromper d'environ de 20%. C'est d'autant plus vrai quand on compare les post-probabilités des scénarios pour les populations alpines, ou pour tous les scénarios faisant intervenir un déclin, une faible différence de valeur est présente par exemple pour la chartreuse, la probabilité du meilleur modèle est de 0.42 contre 0.29 et 0.27 pour les deux autres modèles ce qui est fortement semblable et a donc pour conséquence un fort taux d'erreur de type 1 avec 64% d'erreur. Il semble aussi que les différentes probabilités de choix de scénario ne soient pas significatives, notamment pour certains groupes des alpes du sud, ou on voit des probabilités non significatives avec des intervalles de confiance chevauchant.

Pour la recherche du moment de split, une approche en duo-population a été réalisée avec DIYABC, pour tous les couples de massifs. Pour limiter le nombre de couples et puisqu'il semble avoir une histoire assez similaire pour chaque massif, un représentant des alpes du nord (groupe Bhsj) et un représentant des alpes du sud (groupe Ver-bell) ont été sélectionné. Un premier résultat assez net, est que pour tous les couples le split a eu lieu entre le LGM et le mid-Holocène. En effet, pour chaque couple, aucun intervalle de confiance ne se chevauche et une moyenne de probabilité de 0.99 (Table sup 2). Concernant les erreurs, les erreurs de type 1 et de type 1 semblent être similairement importantes étant aux alentours de 25% (moins que les monos-populations) sauf pour le couple Alpes_sud/Auvergne avec une erreur très faible de 0.02 et 0.0057. Les temps de split estimé pour chaque couple sont recensé dans la table 2 accompagnée de leur intervalle de confiance. En fonction du couple ce temps peut varier de 12 900 ans avant maintenant jusqu'à il y a 1050 ans. Avec cette étude, il semble apparaître plusieurs phases de split. Les couples ayant les splits les plus anciens semblent concerner le massif central avec le groupe AC/Auvergne et la séparation du groupe AC en Ardèche et Cévennes ayant eu lieu à un moment similaire, il y a environ 13 000 ans.

Au même moment, une séparation semble avoir lieu entre les Alpes du Nord et les Pyrénées il y a 11 600 ans. Une seconde phase de split a eu lieu aux alentours de 7000 ans avec un split entre Alpes du sud/Pyrénées ainsi que Pyrénées/AC. Une troisième phase aux alentours des 3000 ans avec un split entre Auvergne/Alpes du sud, l’Auvergne/Pyrénées, Alpes du sud/AC, Alpes du nord/AC ainsi que Alpes du Nord /Auvergne. Enfin une dernière phase de split plus récent semblant concerné plus précisément les alpes, et son éclatement en 3 groupes il y à environ 1000 ans avant maintenant (Table 2). En regardant plus précisément ces temps de split avec les différents intervalles de confiance, il semble alors difficile d’établir un ordre simple d’apparition de population. Cette difficulté à établir cet ordre peut être causée par la présence d’un flux de gène entre ses différents couples, ce qui est testé avec la méthode DILS (Table sup 3). Avec celle-ci, et pour les mêmes couples de population, le modèle choisi parmi les 4 testés est celui représentant une isolation avec une migration continue, modèle choisi avec des valeurs de probabilité allant de 0.64 à 0.93. Les taux de migrations estimés peuvent varier de 0.31% jusqu’à quasi 7%, correspondant à la proportion de la population source se déplaçant. De plus, les taux de migration sont symétriques pour chaque couple (flux de la pop 1 à 2 et de la pop 2 à 1). Cependant les flux de gènes sont hétérogènes: certains locus sont imperméables à la migration, ce sont des loci dits ‘barrière’ potentiellement impliqués dans l’isolement reproducteur et la spéciation (Butlin and Smadja, 2018)

Couples	Années avant maintenant	IC
Auw/AC	12800	[4 510-19 000]
verbell/pyr	7260	[3 260-10 400]
Pyr/AC	7210	[2 840-11 000]
Auw/Pyr	2980	[679-16 200]
Auw/verbell	2580	[1 020-4 300]
verbell/AC	2130	[847-3 340]
bhsj/char	1350	[446-2 170]
verbell/char	1050	[320-1 750]
verbell/bhsj	1120	[352-1 800]
bhsj/AC	3880	[1 420-5 610]
bhsj_pyr	11600	[6 060-14 800]
bhsj/auv	3350	[1 350-5 250]
Ard/Cev	12900	[9 900-42 900]

Table 2 : Estimation du temps de split pour les différents couples de massifs

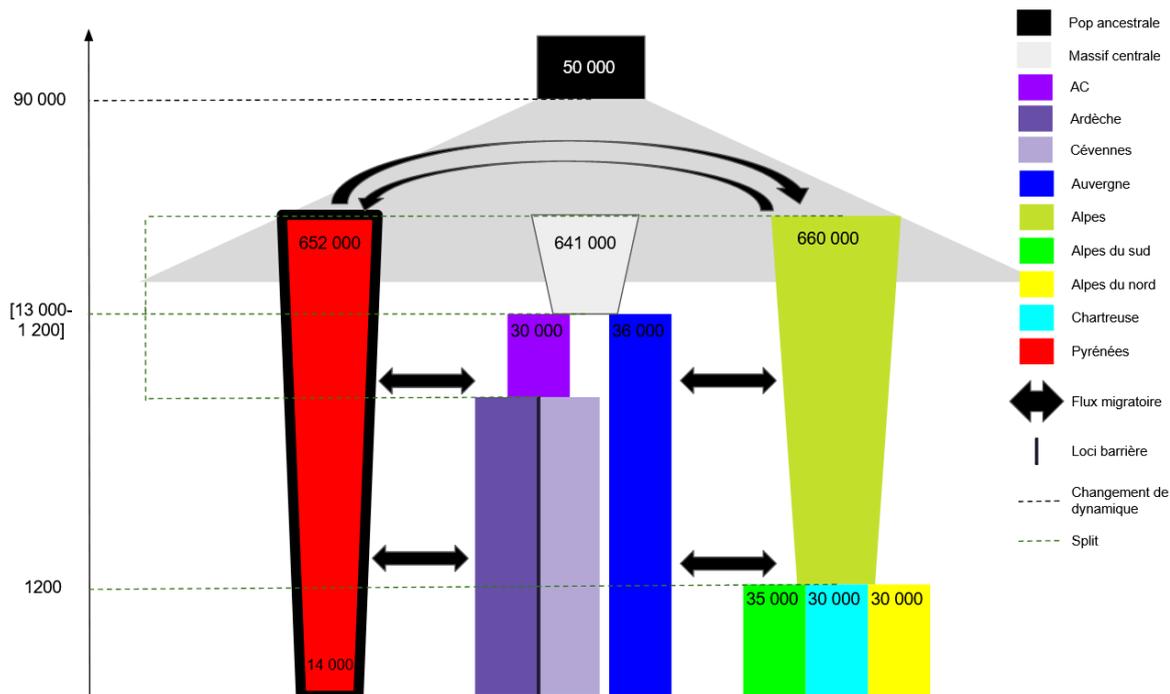


Figure 4: Scénario démographique final

Les chiffres représentent les estimations de taille efficace (N_e) aux différents temps depuis le LIG jusqu'au présent. La présence de loci barrières (imperméable au flux de gène) dans les Pyrénées est représentée par un trait épais autour de cette population..

En regroupant l'ensemble des informations et des résultats obtenus, une figure bilan a été réalisée afin de reconstruire l'histoire évolutive de *Parnassius Apollo* dans les principaux massifs français (Figure 4). La dynamique globale semble bien correspondre à une population ancestrale de taille proche de 50 000 qui a subi une expansion il y a environ 90 000 ans puis une phase de déclin continue démarrant il y a 13 000. Concernant les données récentes, avec l'approche en mono-population grâce à DIYABC, il semble que chaque massif ait une taille récente similaire aux alentours de 30 000 individus bien qu'un peu plus faible pour les Pyrénées mais la différence n'est pas significative (Fig sup 4). Un autre résultat important est le scénario retrouvé avec la méthode DILS, qui est une divergence avec migration. Les flux de gènes retrouvés semblent tout de même assez faibles avec seulement quelques individus échangés (Table sup 3). Sachant que l'ensemble des données numériques a été obtenu à l'aide DIYABC, une méthode ne prenant pas en compte ce potentiel taux de migration pose alors un fort questionnement.

Discussion

Une population structurée par massif

La diversité ainsi que la structure génétique estimées dans cette étude semble dessiner un schéma assez clair pouvant faire office d'une vision d'ensemble. En effet, au niveau de la diversité génétique, les populations du massif central semblent éloignées génétiquement. En effet le F_{st} population spécifique est nettement plus élevé en Auvergne, Ardèche et Cévenne que pour toutes les autres régions.

En regardant un peu plus précisément ce massif, l'Auvergne semble être le groupe le plus éloigné génétiquement (Valeur maximal pour les pairwise F_{ST}), confirmé par l'EEMS montrant des barrières assez fortes entourant les populations auvergnates. Cependant la différenciation au sein de ce massif semble être particulièrement expliquée par l'isolation par la distance et non par la dérive. Cet aspect pourrait donc signifier que les populations auvergnates sont assez similaires mais n'ont pas les capacités pour se déplacer vers leurs plus proches voisins.

Pour l'Ardèche et les Cévennes, la structure est plus complexe car très métissée par les alpes et l'auvergne. On trouve avec les haplotypes que le groupe AC est plus proche de l'Auvergne alors qu'en regardant avec les SNPs et notamment les pairwise F_{ST} le groupe AC est significativement plus proche des groupes alpins. De plus, les populations de cette zone semblent en effet bien isolées, visibles notamment avec EEMS pouvant donc expliquer la faible diversité génétique et donc la présence de dérive.

Inversement, les populations alpines semblent être dans de bien meilleures conditions. En effet, il semble avoir une diversité génétique plutôt importante, étant les seules populations ayant un D_{Tajima} plutôt négatif bien que pas significatif signe d'une population en potentiel expansion. Aussi la différenciation des Alpes en trois sous-groupes pourrait être expliquée par la présence barrière notamment entre les alpes du sud et du nord avec la vallée du Grésivaudan semblant trop grande pour être franchis. Une nouvelle campagne d'échantillonnage un peu plus précise dans cette zone comme en Maurienne pourrait confirmer cette idée.

Enfin les Pyrénées semblent similaires aux alpes, notamment au niveau des pairwise F_{ST} ce qui peut être expliqué par la taille similaire des deux zones. Au niveau de la différenciation, elle semble être assez bien expliquée par de l'isolation par la distance bien qu'il en ressort une zone favorable à la migration. Ceci peut s'expliquer par les peu de points présents en Pyrénées et étant particulièrement espacés. Il faudrait alors plus d'échantillons pour avoir un gradient plus précis d'est en ouest.

Une histoire évolutive commune

Globalement pour chacun des massifs on retrouve une histoire similaire avec une explosion démographique semblant liée à la fin du LIG considéré comme une des périodes les moins favorables pour ce papillon (Kebaïli et al., 2022) avec notamment des températures plus chaudes que de nos jours (Wilcox et al., 2020). La fin de cette période correspond à une phase de température qui se refroidit graduellement pour arriver au dernier maximum glaciaire (LGM). La baisse des températures peut alors être plus propice à l'établissement de cette espèce. Cependant cette baisse de température peut aussi être un problème pour ce papillon, notamment à cause d'un recouvrement des massifs européens par de la calotte glaciaire bloquant ainsi l'accès au massif. Cette idée pourrait alors expliquer la phase de faible décroissance au moment du LGM après l'expansion des populations. Vient ensuite l'holocène qui est accompagné d'une période de réchauffement qui semble bien conduire à un déclin de toutes les populations, notamment en obligeant les différentes espèces déjà établies à s'adapter en montant en altitude (Réchauffement qui va libérer les sommets de la glace). Cette idée est confirmée à la fois par le stairway plot et par DIYABC, permettant donc de trancher sur les deux scénarios hypothèse, présenté en début d'étude, avec ici l'évidence d'une fragmentation pendant la contraction de la population ancestrale conduisant à l'isolement entre chacun des massifs. Aussi, il s'avère particulièrement difficile avec nos résultats d'estimer quelles populations ont divergé en premier. Ceci peut être expliqué par le déclin en lui-même, étant continue au cours du temps.

Une autre question importante à se poser est la présence de refuges, de nombreuses théories ici s'affrontent avec l'existence d'un potentiel refuge dans l'Auvergne ou dans le massif central (Kebaïli et al., 2022). Dans notre étude, le seul indice allant dans ce sens est l'aspect très métissés du groupe Ardèche/Cévennes qui aurait pu être une bonne zone candidate. Cependant par définition, une zone refuge est synonyme d'une bonne diversité génétique (Origine de la population ancestrale). Or dans cette étude, on trouve que le Massif Central et notamment l'Ardèche et les Cévennes ont le plus faible H_e contredisant alors cette idée de refuge. De plus, sachant que pour tous les massifs, la même histoire est retrouvée avec plusieurs méthodes différentes ne va vraiment pas dans l'idée d'un refuge mais plus vers l'idée d'une colonisation par une population source. Cette idée est supportée par Todisco & al, qui à l'aide d'une étude sur de l'ADN mitochondrial, retrouve une colonisation de P.apollo de la France depuis l'Asie. L'ADN mitochondrial est notamment pertinente pour des événements de divergence anciens et en présence de large structure génétique à de grandes échelles spatiales (Ehl et al., 2021; Schmitt, 2007). De plus, l'ADNmt semble réagir différemment de l'ADN nucléaire par rapport aux fluctuations démographique obligeant donc

l'utilisation d'autre marqueur pour l'estimation d'histoire démographique complexe de population n'étant pas en équilibre. (Després et al., 2019; Ebdon et al., 2021; Kebaïli et al., 2022; Schoville and Roderick, 2009). Dans ce scénario, Todisco & al décrit une première colonisation des alpes qui aurait agit comme un berceau de colonisation secondaire pour le reste de la france. Ceci semble plus cohérent avec la diversité génétique calculée dans cette étude, avec les populations alpines ayant le He le plus important. En sachant qu'avec notre étude qui n'est pas focalisée sur de l'adn mitochondrial, on retrouve cette idée de colonisation, ce scénario devient alors un peu plus crédible. Enfin, un résultat surprenant est l'idée que les populations ne sont pas strictement isolées et continuent à communiquer entre elles, résultat divergeant totalement avec Kebaïli & al. En effet avec la méthode DILS, des taux de migration faible mais strictement supérieurs à 0 sont retrouvés entre tous les massifs allant donc dans le sens d'une connexion entre les différents massifs est toujours présent. Ceci pourrait expliquer pourquoi il est difficile d'estimer précisément un ordre d'apparition des populations, avec un déclin continu avec des populations toujours en contact. Néanmoins ces résultats nécessitent d'être approfondis par d'autres méthodes car des incohérence sont présentes avec DILS.

Limites et perspectives

En effet, un fait marquant avec ces résultats est la présence de taux de migration totalement symétrique entre chaque population ce qui est inattendu quand on compare les populations de d'Auvergne, peu nombreuses et avec des indices génétiques plutôt négatifs, avec les populations alpines, beaucoup plus nombreuses. Classiquement, l'attente serait d'avoir un flux beaucoup plus important venant des Alpes, ayant des meilleures conditions pour assurer le rôle de source.

De plus, sachant que DIYABC et le stairway plot ne prennent pas en compte des taux de migrations, et que DILS en trouve, cela laisse donc supposer que les tailles inférées sont probablement biaisées, expliquant pourquoi on trouve une taille récente similaire pour toutes les populations. Il serait plus attendu que la taille des populations du massif centrale (groupe Auvergne et AC) d'avoir une taille nettement plus faible que les alpes et les pyrénées, ce qui est visible dans l'approche en duo-pop avec DIYABC et DILS.

Cependant l'inférence de taille de population par DILS semble bien trop grande sans aucune cohérence avec les méthodes précédentes et pose donc un problème de confiance.

Aussi cette analyse peut être fortement approfondie, et le scénario évolutif définitivement tranché par un échantillonnage plus important à la fois en France et potentiellement en

Europe. En effet, la zone géographique ayant des individus ne sont pas présente dans cette étude et pourrait conforter l'idée d'une colonisation ou non, comme l'Ariège en France ou alors le nord de l'Italie.

Un autre axe d'étude est aussi possible, avec notamment quelques locis barrière retrouvées avec Dils pour tous les couples comprenant les Pyrénées. Ces locis pourraient donc être un indice d'adaptation locale dans cette zone. Par exemple, la spécificité plante-hôte peut jouer un rôle avec notamment des différences entre chaque massif (Deschamps-Cottin et al., 1997).

Conclusion

En conclusion à l'aide d'une approche en deux temps, avec en premier lieu une étude sur la structure génétique des population actuelles ainsi que la réalisation d'inférence démographique, cette étude à permis de faire un constat actuelle, comme par exemple les populations d'Auvergne ayant des indicateur de diversité alarmant, ainsi que l'évidence d'indice sur le passé mystérieux de cette espèce de papillon emblématique. En effet, cette étude est en faveur d'un scénario, retrouvé dans la plus récente étude sur cette espèce en date, c'est à dire une forte croissance démographique pendant la dernière période de refroidissement (entre le LIG et le LGM) suivi d'un déclin brutal pendant l'holocène (période de réchauffement). Cette phase de déclin aurait donc entraîné la fragmentation de la population ancestrale dans chaque massif. Aussi cette étude à permis de mettre en lumière de nouveaux indices sur l'origine de cette population en France, avec l'idée d'une colonisation depuis l'extérieur retrouvée dans la littérature. Néanmoins un aspect étonnant de cette étude est le fait de trouver une divergence avec flux de gènes entre les différents massifs bien qu'une forte structure soit présente pour chacune des populations. Cette information difficilement explicable mériterait d'être un nouvel axe de recherche permettant de mieux comprendre l'état de ces populations et ainsi une meilleure approche conservative face aux enjeux du réchauffement climatique.

Référence

- Booker, T.R., Jackson, B.C., Keightley, P.D., 2017. Detecting positive selection in the genome. *BMC Biol.* 15, 98. <https://doi.org/10.1186/s12915-017-0434-y>
- Butlin, R.K., Smađja, C.M., 2018. Coupling, Reinforcement, and Speciation. *Am. Nat.* 191, 155–172. <https://doi.org/10.1086/695136>
- Cornuet, J.-M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., Marin, J.-M., 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. <https://doi.org/10.1093/bioinformatics/btt763>
- Cornuet, J.-M., Ravigné, V., Estoup, A., 2010. Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *BMC Bioinformatics* 11, 401. <https://doi.org/10.1186/1471-2105-11-401>
- Deschamps-Cottin, M., Roux, M., Descimon, H., 1997. Valeur trophique des plantes nourricières et préférence de ponte chez *Parnassius apollo* L. (Lepidoptera, Papilionidae). *Comptes Rendus Académie Sci. - Ser. III - Sci. Vie* 320, 399–406. [https://doi.org/10.1016/S0764-4469\(97\)85028-7](https://doi.org/10.1016/S0764-4469(97)85028-7)
- Després, L., Henniaux, C., Rioux, D., Capblancq, T., Zupan, S., Ćelik, T., Sielezniew, M., Bonato, L., Ficetola, G.F., 2019. Inferring the biogeography and demographic history of an endangered butterfly in Europe from multilocus markers. *Biol. J. Linn. Soc.* 126, 95–113. <https://doi.org/10.1093/biolinnean/bly160>
- Dray, S., 2007. The ade4 package-II: Two-table and K-table methods. *R News* 47–52.
- Ebdon, S., Laetsch, D.R., Dapporto, L., Hayward, A., Ritchie, M.G., Dincă, V., Vila, R., Lohse, K., 2021. The Pleistocene species pump past its prime: Evidence from European butterfly sister species. *Mol. Ecol.* 30, 3575–3589. <https://doi.org/10.1111/mec.15981>
- Ehl, Stefan, Ehl, Sarah, Kramp, K., Schmitt, T., 2021. Interglacials are driving speciation and intraspecific differentiation in the cold-adapted butterfly species group *Boloria pales* / *napaea* (Nymphalidae). *J. Biogeogr.* 48, 134–146. <https://doi.org/10.1111/jbi.13988>
- Fraïsse, C., Popovic, I., Mazoyer, C., Spataro, B., Delmotte, S., Romiguié, J., Loire, É., Simon, A., Galtier, N., Duret, L., Bierne, N., Vekemans, X., Roux, C., 2020. DILS: Demographic Inferences with Linked Selection by using ABC (preprint). *Evolutionary Biology*. <https://doi.org/10.1101/2020.06.15.151597>
- Goudet, J., 2005. hierfstat, a package for r to compute and test hierarchical F-statistics. *Mol. Ecol. Notes* 5, 184–186. <https://doi.org/10.1111/j.1471-8286.2004.00828.x>
- Hewitt, G., 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol. J. Linn. Soc.* 58, 247–276. <https://doi.org/10.1006/bjll.1996.0035>
- Hewitt, G.M., 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 359, 183–195. <https://doi.org/10.1098/rstb.2003.1388>
- Hijmans, R.J., 2021. Introduction to the “geosphere” package.
- Kebaili, C., Sherpa, S., Rioux, D., Després, L., 2022. Demographic inferences and climatic niche modelling shed light on the evolutionary history of the emblematic cold-adapted Apollo butterfly at regional scale. *Mol. Ecol.* 31, 448–466. <https://doi.org/10.1111/mec.16244>
- Keightley, P.D., Pinharanda, A., Ness, R.W., Simpson, F., Dasmahapatra, K.K., Mallet, J., Davey, J.W., Jiggins, C.D., 2015. Estimation of the Spontaneous Mutation Rate in *Heliconius melpomene*. *Mol. Biol. Evol.* 32, 239–243. <https://doi.org/10.1093/molbev/msu302>
- Kukla, G.J., Bender, M.L., de Beaulieu, J.-L., Bond, G., Broecker, W.S., Cleveringa, P., Gavin, J.E., Herbert, T.D., Imbrie, J., Jouzel, J., Keigwin, L.D., Knudsen, K.-L., McManus, J.F., Merkt, J., Muhs, D.R., Müller, H., Poore, R.Z., Porter, S.C., Seret, G., Shackleton, N.J., Turner, C., Tzedakis, P.C., Winograd, I.J., 2002. Last Interglacial Climates. *Quat. Res.* 58, 2–13. <https://doi.org/10.1006/qres.2001.2316>
- Liu, X., Fu, Y.-X., 2015. Exploring population size changes using SNP frequency spectra. *Nat. Genet.* 47, 555–559. <https://doi.org/10.1038/ng.3254>
- Malinsky, M., Trucchi, E., Lawson, D.J., Falush, D., 2018. RADpainter and fineRADstructure: Population Inference from RADseq Data. *Mol. Biol. Evol.* 35, 1284–1290. <https://doi.org/10.1093/molbev/msy023>
- Nazareno, A.G., Bemmels, J.B., Dick, C.W., Lohmann, L.G., 2017. Minimum sample sizes for population genomics: an empirical study from an Amazonian plant species. *Mol. Ecol. Resour.* 17, 1136–1147. <https://doi.org/10.1111/1755-0998.12654>
- Ometto, L., Li, M., Bresadola, L., Barbaro, E., Neteler, M., Varotto, C., 2015. Demographic History, Population Structure, and Local Adaptation in Alpine Populations of *Cardamine impatiens* and *Cardamine resedifolia*. *PLOS ONE* 10, e0125199. <https://doi.org/10.1371/journal.pone.0125199>
- Petkova, D., Novembre, J., Stephens, M., 2016. Visualizing spatial population structure with estimated effective migration surfaces. *Nat. Genet.* 48, 94–100. <https://doi.org/10.1038/ng.3464>
- Podsiadlowski, L., Tunström, K., Espeland, M., Wheat, C.W., 2021. The Genome Assembly and Annotation of the Apollo Butterfly *Parnassius apollo*, a Flagship Species for Conservation Biology. *Genome Biol. Evol.* 13, evab122. <https://doi.org/10.1093/gbe/evab122>
- Schmitt, T., 2007. Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Front. Zool.* 4, 11. <https://doi.org/10.1186/1742-9994-4-11>
- Schoville, S.D., Roderick, G.K., 2009. Alpine biogeography of Parnassian butterflies during Quaternary climate cycles in North America. *Mol. Ecol.* 18, 3471–3485. <https://doi.org/10.1111/j.1365-294X.2009.04287.x>
- Schoville, Sean D., Stuckey, M., Roderick, G.K., 2011. Pleistocene origin and population history of a neoendemic alpine butterfly: POPULATION HISTORY OF *COLIAS BEHRII*. *Mol. Ecol.* 20, 1233–1247. <https://doi.org/10.1111/j.1365-294X.2011.05003.x>
- Strandberg, G., Brandefelt, J., Kjellström, E., Smith, B., 2011. High-resolution regional simulation of last glacial maximum climate in Europe. *Tellus Dyn. Meteorol. Oceanogr.* 63, 107–125. <https://doi.org/10.1111/j.1600-0870.2010.00485.x>
- Taberlet, P., Fumagalli, L., Wust-Saucy, A., Cosson, J., 1998. Comparative phylogeography and postglacial colonization routes in Europe. *Mol. Ecol.* 7, 453–464. <https://doi.org/10.1046/j.1365-294x.1998.00289.x>
- Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595. <https://doi.org/10.1093/genetics/123.3.585>
- Teixeira, H., Salmons, J., Arredondo, A., Mourato, B., Manzi, S., Rakotondravony, R., Mazet, O., Chikhi, L., Metzger, J., Radespiel, U., 2021. Impact of model assumptions on demographic inferences: the case study of two sympatric mouse lemurs in northwestern Madagascar. *BMC Ecol. Evol.* 21, 197. <https://doi.org/10.1186/s12862-021-01929-z>
- Todisco, V., Gratton, P., Cesaroni, D., Sbordoni, V., 2010. Phylogeography of *Parnassius apollo*: hints on taxonomy and conservation of a vulnerable glacial butterfly invader: PHYLOGEOGRAPHY OF *PARNASSIUS APOLLO*. *Biol. J. Linn. Soc.* 101, 169–183. <https://doi.org/10.1111/j.1095-8312.2010.01476.x>
- Wilcox, P.S., Honiat, C., Trüssel, M., Edwards, R.L., Spötl, C., 2020. Exceptional warmth and climate instability occurred in the European Alps during the Last Interglacial period. *Commun. Earth Environ.* 1, 57. <https://doi.org/10.1038/s43247-020-00063-w>

Annexe

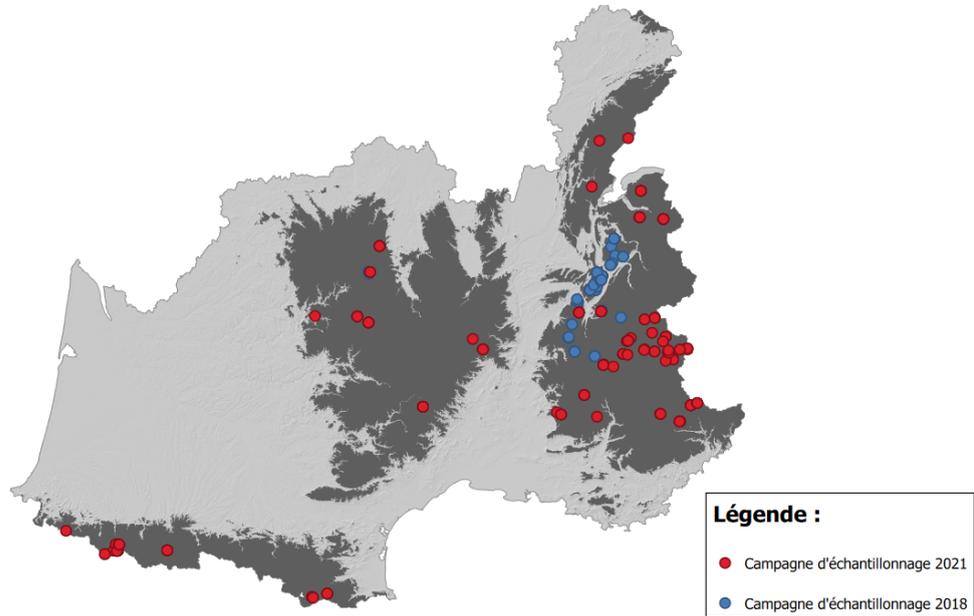


Figure sup 1 : Carte d'échantillonnage des 283 individus utilisé dans cette études
Dans cette figure, chaque point représente un individu issu soit d'une campagne d'échantillonnage en 2018 (bleue) soit en 2021 (rouge).

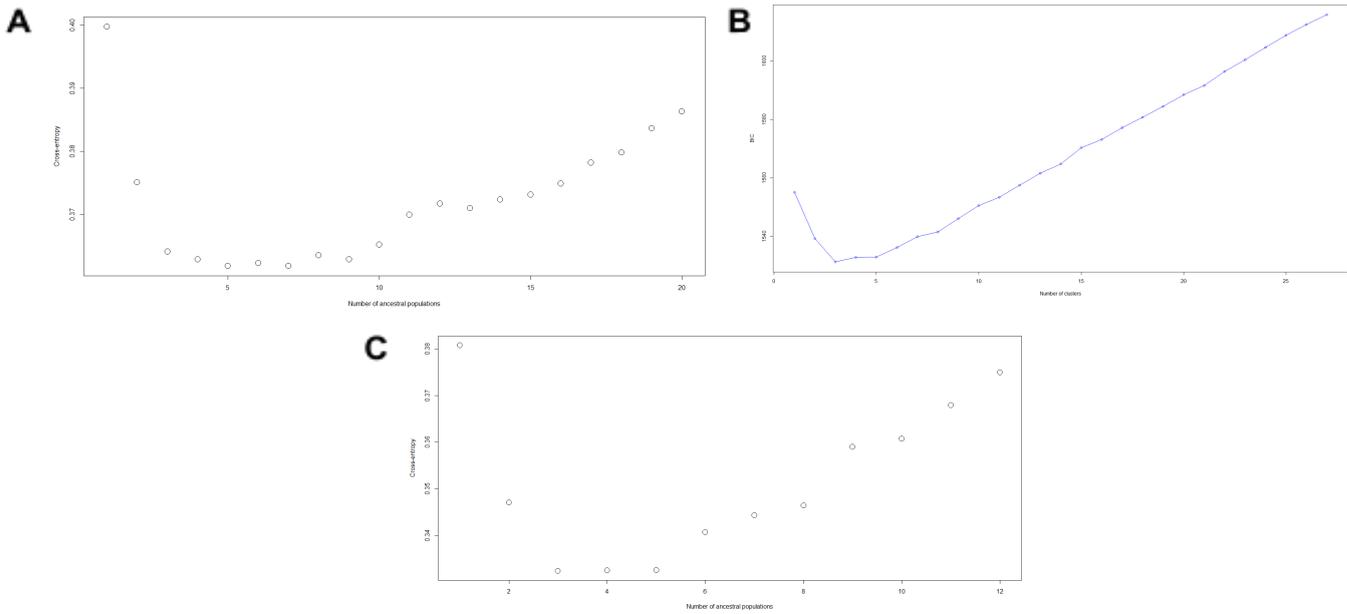


Figure sup 2 : Détermination du K optimale par 3 méthode différentes

K optimale correspondant à la valeur de cross-entropy la plus faible déterminé avec la méthode snmf (A), avec DAPC (B) ainsi qu'avec Admixture (C)

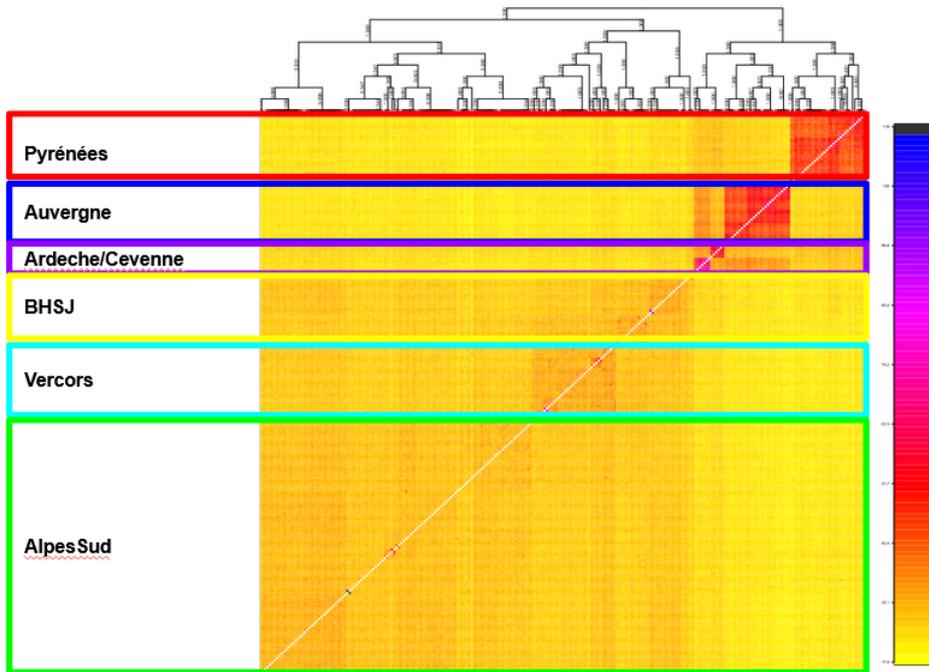


Figure sup 3 : Structure génétique des populations de *P.apollo* pour tous les massifs à l'aide des haplotypes

Cette figure évalue l'apparementement deux à deux pour chaque paire d'individus, ainsi la diagonale blanche représente ici la comparaison de deux individus identiques.

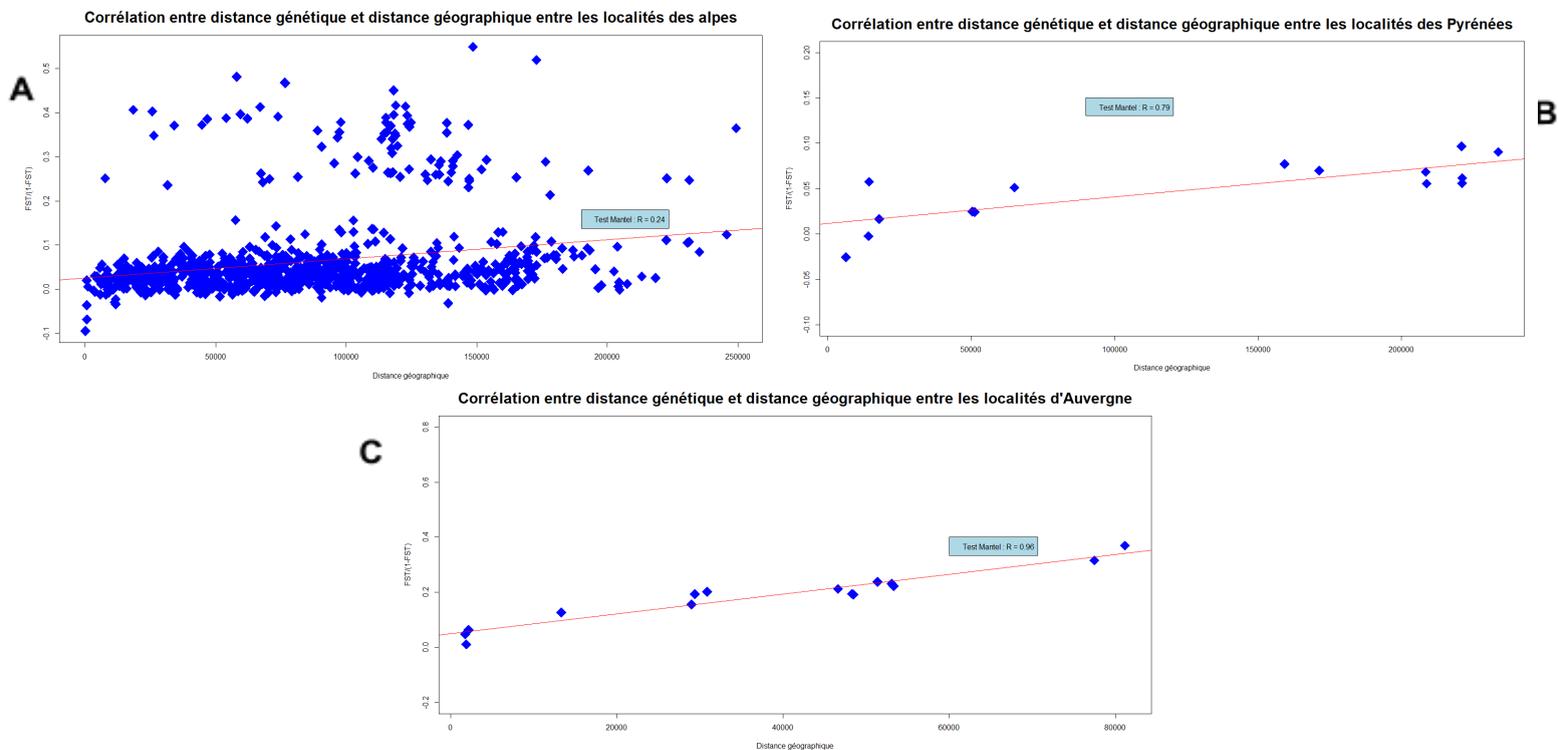


Figure sup 3 : Isolation par la distance dans les massifs des Alpes, de l'Auvergne et des Pyrénées
Corrélation entre la distance génétique corrigée $FST/(1-FST)$ en fonction de la distance géographique pour les Alpes (A), les Pyrénées (B)

et pour l'Auvergne (C). Pour chaque corrélation, un test de Mantel à été réalisé pour savoir si la corrélation est significative ou non.

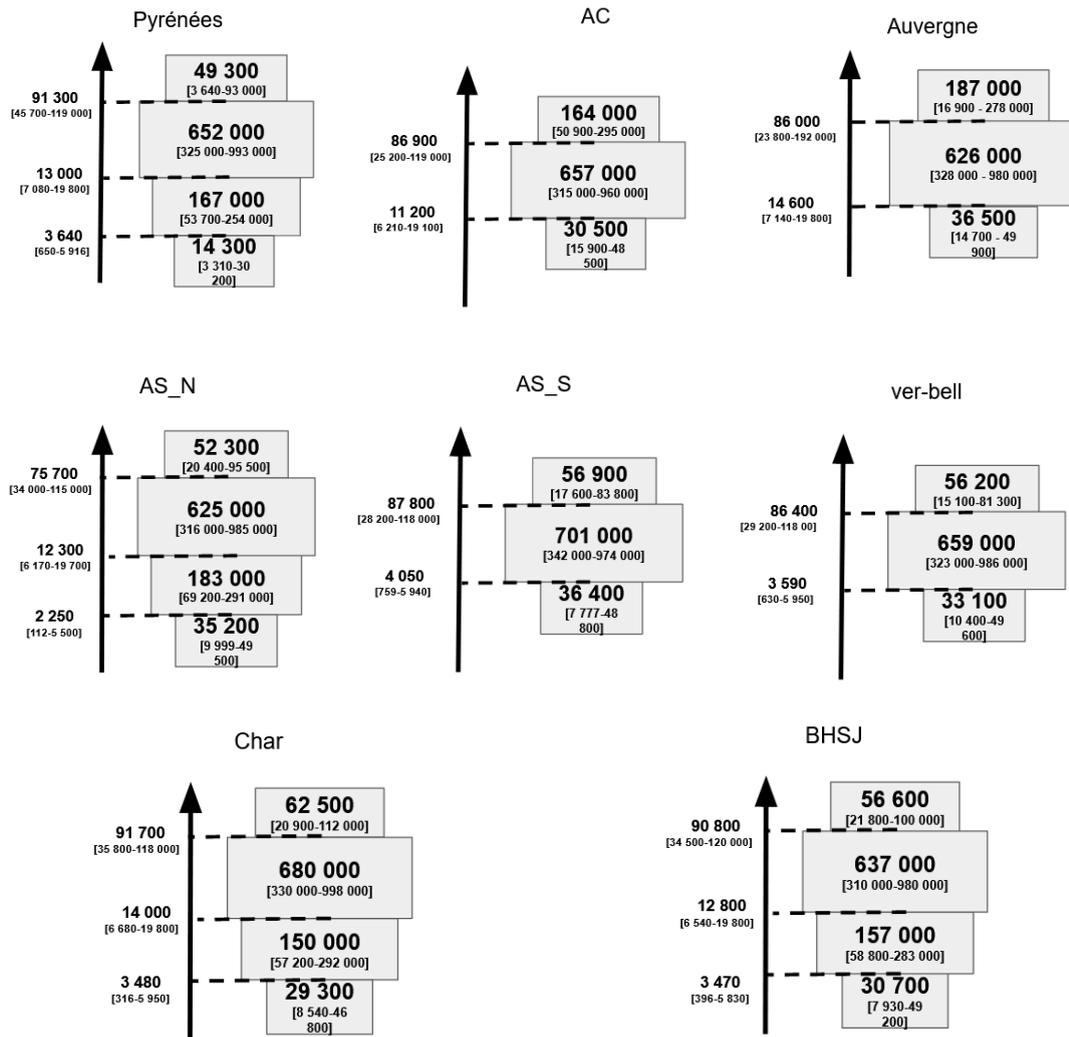


Figure sup 4 : Meilleur scénario démographique avec un modèle en 1 dimension

Représentation du meilleurs scénario choisi avec l'approche mono-populationnel avec DIYABC pour chaque population

Région	Zone	Population	Freq	Ho	He	Fis	fis_95PCI	Fst	fst_95PCI	TajimaD	TajimaD_95PCI
Alpes			192	0.094	0.135	0.175		0.188			
	Alpe du sud		81	0.091	0.133	0.191		0.191		-0.492	[-1.64/1.353]
		GSP	5	0.090	0.133	0.229	[0.206-0.254]	0.126	[0.099-0.157]		
		LUR	1	0.087	NA	NA	[NA-NA]	0.392	[0.331-0.449]		
		VEN	5	0.090	0.137	0.239	[0.213-0.265]	0.115	[0.079-0.15]		
		AGN	2	0.096	0.134	0.116	[0.076-0.156]	0.194	[0.147-0.24]		
		ARV	2	0.095	0.126	0.101	[0.064-0.142]	0.222	[0.173-0.267]		
		BRA	3	0.093	0.139	0.190	[0.158-0.222]	0.136	[0.101-0.173]		
		BUR	5	0.099	0.139	0.186	[0.164-0.209]	0.091	[0.061-0.122]		
		CEI	3	0.099	0.135	0.152	[0.122-0.182]	0.132	[0.094-0.175]		
		CER	4	0.086	0.137	0.251	[0.224-0.277]	0.126	[0.09-0.161]		
		CHL	4	0.099	0.137	0.164	[0.14-0.191]	0.116	[0.081-0.151]		
		DOR	3	0.085	0.134	0.231	[0.201-0.263]	0.161	[0.12-0.203]		
		ESC	3	0.084	0.127	0.202	[0.171-0.238]	0.209	[0.167-0.253]		
		HRE	1	0.091	NA	NA	[NA-NA]	0.359	[0.294-0.42]		
		LZO	5	0.093	0.134	0.204	[0.183-0.229]	0.123	[0.093-0.155]		
		PAR	1	0.087	NA	NA	[NA-NA]	0.393	[0.334-0.454]		
		RAB	1	0.092	NA	NA	[NA-NA]	0.330	[0.266-0.401]		
		RRO	2	0.089	0.127	0.145	[0.107-0.185]	0.233	[0.186-0.281]		
		STC	3	0.088	0.129	0.200	[0.171-0.232]	0.184	[0.141-0.223]		
		VDV	5	0.093	0.135	0.214	[0.192-0.239]	0.113	[0.08-0.146]		
		VIS	5	0.084	0.129	0.240	[0.216-0.266]	0.166	[0.135-0.196]		
		BLA	5	0.092	0.136	0.213	[0.19-0.238]	0.118	[0.086-0.147]		
		DAL	6	0.085	0.125	0.232	[0.208-0.255]	0.170	[0.139-0.2]		
		ISO	5	0.092	0.133	0.212	[0.189-0.238]	0.127	[0.096-0.158]		
		LOV	2	0.092	0.127	0.094	[0.051-0.135]	0.251	[0.205-0.299]		
	Vercors		28	0.101	0.134	0.161		0.124		-0.309	[-1.654/1.647]
		COR	4	0.098	0.131	0.162	[0.135-0.189]	0.154	[0.115-0.193]		
		PIC	5	0.107	0.138	0.146	[0.122-0.17]	0.094	[0.061-0.13]		
		MOL	5	0.103	0.140	0.171	[0.148-0.195]	0.089	[0.057-0.123]		
		TUS	5	0.105	0.141	0.167	[0.145-0.192]	0.081	[0.047-0.114]		
		LUS	5	0.098	0.138	0.192	[0.169-0.216]	0.105	[0.075-0.139]		
		LEV	4	0.093	0.117	0.130	[0.097-0.166]	0.218	[0.177-0.253]		
	Belledonne		8	0.096	0.135	0.186		0.130		-0.14	[-1.692/1.582]
		BE1	5	0.100	0.139	0.188	[0.164-0.213]	0.091	[0.057-0.124]		
		CHR	3	0.091	0.131	0.183	[0.154-0.214]	0.169	[0.129-0.208]		
	Chartreuse		36	0.103	0.140	0.150		0.110		-0.098	[-1.489/1.794]
		COQ	3	0.105	0.144	0.141	[0.114-0.173]	0.100	[0.058-0.143]		
		SCI	5	0.099	0.142	0.202	[0.179-0.225]	0.079	[0.044-0.113]		
		FIT	4	0.105	0.146	0.170	[0.146-0.196]	0.069	[0.034-0.105]		
		SOM	5	0.102	0.138	0.170	[0.147-0.195]	0.101	[0.069-0.137]		
		SUI	2	0.115	0.137	0.039	[0.002-0.079]	0.127	[0.075-0.179]		
		RAC	4	0.099	0.140	0.186	[0.159-0.215]	0.103	[0.067-0.139]		
		GRA	6	0.105	0.138	0.152	[0.133-0.175]	0.092	[0.06-0.127]		
		OUT	5	0.102	0.137	0.162	[0.136-0.186]	0.114	[0.079-0.15]		
		PRA_C	2	0.095	0.138	0.129	[0.085-0.172]	0.202	[0.155-0.25]		
	Bauges		21	0.104	0.136	0.141		0.153		-0.221*	[-1.558/1.703]*
		MON	1	0.094	NA	NA	[NA-NA]	0.349	[0.281-0.416]		
		SEM	5	0.107	0.133	0.125	[0.1-0.151]	0.126	[0.091-0.161]		
		COC	5	0.110	0.135	0.109	[0.086-0.131]	0.112	[0.078-0.148]		

		ARM	5	0.104	0.141	0.185	[0.163-0.208]	0.068	[0.033-0.107]		
		CRO	5	0.104	0.136	0.144	[0.12-0.167]	0.109	[0.074-0.141]		
	Haute-savoie		15	0.089	0.130	0.219		0.145		*	*
		AYE	5	0.086	0.127	0.225	[0.199-0.253]	0.166	[0.133-0.2]		
		BGY	5	0.093	0.134	0.206	[0.182-0.231]	0.121	[0.092-0.153]		
		HIR	5	0.088	0.130	0.225	[0.202-0.251]	0.149	[0.117-0.181]		
	Jura		3	0.078				0.461		*	*
		MDO	1	0.072	NA	NA	[NA-NA]	0.499	[0.441-0.553]		
		VUI	1	0.078	NA	NA	[NA-NA]	0.463	[0.41-0.522]		
		PRE	1	0.083	NA	NA	[NA-NA]	0.420	[0.358-0.482]		
Massif Centrale			41	0.066	0.087	0.159		0.424			
	Ardèche		7							0.44	[-1.481/2.163]
		BOC	7	0.060	0.084	0.214	[0.185-0.247]	0.445	[0.414-0.476]		
	Cévennes		6							0.404	[-1.486/2.006]
		MEJ	6	0.078	0.106	0.187	[0.162-0.217]	0.304	[0.268-0.338]		
	Auvergne		28	0.061	0.072	0.077		0.524		0.623	[-1.452/2.646]
		PRA	5	0.062	0.071	0.077	[0.041-0.114]	0.522	[0.491-0.555]		
		RED	5	0.070	0.080	0.072	[0.038-0.104]	0.465	[0.433-0.501]		
		SAN	7	0.062	0.071	0.089	[0.059-0.122]	0.519	[0.486-0.553]		
		PDD	5	0.051	0.059	0.097	[0.06-0.139]	0.592	[0.561-0.624]		
		CHA	3	0.060	0.073	0.066	[0.019-0.116]	0.522	[0.487-0.56]		
		PDP	3	0.061	0.076	0.061	[0.015-0.107]	0.525	[0.488-0.563]		
Pyrénées			35	0.083	0.121	0.177		0.244		0.242	[-1.378/2.219]
	Pyrénées de l'ouest		24	0.083	0.121	0.144		0.265			
		ARR	4	0.084	0.125	0.171	[0.136-0.208]	0.241	[0.204-0.283]		
		SOU	4	0.080	0.121	0.197	[0.163-0.234]	0.255	[0.212-0.301]		
		SAG	4	0.078	NA	NA	[NA-NA]	NA	[0.395-0.513]		
		SOM	3	0.083	0.112	0.105	[0.058-0.151]	0.309	[0.262-0.356]		
		SOQ	1	0.083	NA	NA	[NA-NA]	NA	[0.353-0.487]		
		AUL	4	0.095	0.125	0.117	[0.089-0.146]	0.217	[0.175-0.262]		
		BIZ	1	0.081	NA	NA	[NA-NA]	NA	[0.372-0.506]		
		POM	2	0.083	0.120	0.130	[0.097-0.164]	0.304	[0.250-0.359]		
		SEO	1	0.084	NA	NA	[NA-NA]	NA	[0.380-0.467]		
	Pyrénées de l'est		11	0.083	0.121	0.211		0.223			
		EYN	6	0.085	0.123	0.214	[0.188-0.241]	0.201	[0.166-0.237]		
		MAN	5	0.081	0.118	0.207	[0.181-0.237]	0.244	[0.211-0.284]		

Table sup1: Diversité génétique

	Posterior probability		Confidence in scenario choice (errors)				
	Value	CI 95%	Type I	Type II with sc1	Type II with sc2	Type II with sc3	Type II with sc4
Demographic history of each population							
Pyrénées							
Sc1 N	0.0000	[0.0000,0.0000]					
Sc2 ExpDR	0.3648	0.3164 - 0.4131					
Sc3 ExpDA	0.2645	[0.2197,0.3093]					
Sc4 Exp2D	0.3707	[0.3226,0.4189]	0.65	0.02	0.36	0.22	
Auvergne							
Sc1 N	0.0000	[0.0000,0.0000]					
Sc2 ExpDR	0.0812	[0.0537,0.1088]					
Sc3 ExpDA	0.7286	[0.6846,0.7725]	0.45	0.04	0.19		0.24
Sc4 Exp2D	0.1902	[0.1512,0.2291]					
Ardeche/cevennes							
Sc1 N	0.0000	[0.0000,0.0000]					
Sc2 ExpDR	0.2705	[0.2262,0.3148]					
Sc3 ExpDA	0.3930	[0.3448,0.4411]	0.11	0.08	0.01		0.02
Sc4 Exp2D	0.3365	[0.2897,0.3833]					
Chartreuse							
Sc1 N	0.0000	[0.0000,0.0000]					
Sc2 ExpDR	0.2980	[0.2523,0.3436]					
Sc3 ExpDA	0.2779	[0.2335,0.3223]					
Sc4 Exp2D	0.4241	[0.3750,0.4732]	0.64	0.03	0.35	0.21	
AS_S							
Sc1 N	0.0000	[0.0000,0.0000]					
Sc2 ExpDR	0.4955	[0.4464,0.5447]	0.57	0.02		0.14	0.35
Sc3 ExpDA	0.1252	[0.0865,0.1638]					
Sc4 Exp2D	0.3793	[0.3307,0.4280]					
ver-bell							
Sc1 N	0.0000	[0.0000,0.0000]					
Sc2 ExpDR	0.4785	[0.4291,0.5280]	0.63	0.05		0.16	0.35
Sc3 ExpDA	0.0804	[0.0381,0.1227]					
Sc4 Exp2D	0.4411	[0.3908,0.4914]					
bhsj							
Sc1 N	0.0000	[0.0000,0.0000]					
Sc2 ExpDR	0.2799	[0.2355,0.3243]					
Sc3 ExpDA	0.3055	[0.2601,0.3510]					
Sc4 Exp2D	0.4146	[0.3661,0.4631]	0.61	0.01	0.35	0.24	
AS_N							
Sc1 N	0.0000	[0.0000,0.0000]					
Sc2 ExpDR	0.4496	[0.3991,0.5002]					
Sc3 ExpDA	0.1006	[0.0565,0.1447]					
Sc4 Exp2D	0.4498	[0.4007,0.4988]	0.6073	0.01	0.37	0.19	
Demographic simulation with 2 populations							
Alps_Auv							
Sc1 S-LIG-LGM	0.0021	[0.0000,0.0047]					
Sc2 S-LGM-midH	0.9979	[0.9953,1.0000]	0.02	0.0057			
Pyr_Auv							
Sc1 S-LIG-LGM	0.0002	[0.0000,0.0004]					
Sc2 S-LGM-midH	0.9998	[0.9996,1.0000]	0.2499	0.2816			
Pyr_AlpsSud							
Sc1 S-LIG-LGM	0.0000	[0.0000,0.0000]					
Sc2 S-LGM-midH	1.0000	[1.0000,1.0000]	0.2498	0.2815			

AlpesNord_AlpesSud											
Sc1 S-LIG-LGM	0.0000	[0.0000,0.0000]									
Sc2 S-LGM-midH	1.0000	[1.0000,1.0000]	0.24	0.23							
Char_AlpesSud											
Sc1 S-LIG-LGM	0.0000	[0.0000,0.0000]									
Sc2 S-LGM-midH	1.0000	[1.0000,1.0000]	0.25	0.28							
AlpesNord_Char											
Sc1 S-LIG-LGM	0.0000	[0.0000,0.0000]									
Sc2 S-LGM-midH	1.0000	[1.0000,1.0000]	0.19	0.25							
Auv_AC											
Sc1 S-LIG-LGM	0.0016	[0.0004,0.0029]									
Sc2 S-LGM-midH	0.9984	[0.9971,0.9996]	0.27	0.24							
Pyr_AC											
Sc1 S-LIG-LGM	0.0001	[0.0000,0.0003]									
Sc2 S-LGM-midH	0.9999	[0.9997,1.0000]	0.20	0.25							
AlpsSud_AC											
Sc1 S-LIG-LGM	0.0000	[0.0000,0.0000]									
Sc2 S-LGM-midH	1.0000	[1.0000,1.0000]	0.25	0.25							
AlpsNord_AC											
Sc1 S-LIG-LGM	0.0001	[0.0000,0.0001]									
Sc2 S-LGM-midH	0.9999	[0.9999,1.0000]	0.44	0.45							
AlpsNord_pyr											
Sc1 S-LIG-LGM	0.0000	[0.0000,0.0000]									
Sc2 S-LGM-midH	1.0000	[1.0000,1.0000]	0.14	0.22							
AlpsNord_auv											
Sc1 S-LIG-LGM	0.0000	[0.0000,0.0000]									
Sc2 S-LGM-midH	1.0000	[1.0000,1.0000]	0.17	0.21							
Ard_cev											
Sc1 S-LIG-LGM	0.0415	[0.0264,0.0566]									
Sc2 S-LGM-midH	0.9585	[0.9434,0.9736]	0.31	0.34							

Table sup 2: Résultat DIYABC: Choix de scénario et probabilité postérieurs

Couple	Model choice	Post Proba	Migration12	IC_12	Locis barrières 12	Migration21	IC_21	Locis barrières 21	Tsplit	Taille 1	Taille 2
Auv_Pyr	IM	0.79	3.30	[2.36-4.40]	232	4.12	[2.89-5.26]	395	106 275	198 567	268 527
Pyr_AlpeSud	IM	0.80	2.80	[1.82-3.75]	223	1.83	[1.16-2.50]	345	186 596	171 298	375 028
AC_Pyr	IM	0.8	3.8	[2.47-5.26]	0	2.32	[1.79-3.03]	0	176 000	67 021	192 793
Auv_AlpeSud	IM	0.8	2.37	[1.84-3.07]	0	1.39	[1.10-1.67]	0	153 561	68 096	453 061
Ac_AlpeSud	IM	0.81	4.85	[3.92-5.81]	0	1.26	[0.93-1.59]	0	155 034	49 034	387 008
AlpeN_AlpeS	IM	0.93	6.06	[4.52-7.59]	0	6.63	[4.91-8.44]	0	202 414	204 277	294 002
Pyr_bhsj	IM	0.84	3.02	[2.35-3.84]	117	1.34	[0.97-1.76]	88	198 671	171 558	251 186
Auv_AC	IM	0.73	2.86	[2.10-3.58]	0	2.22	[1.42-2.99]	0	22 719	581 359	119 310
Auv_bhsj	IM	0.78	3.39	[2.33-4.55]	0	1.78	[1.33-2.28]	0	106 060	291 000	342 000
Bhsj_Ac	IM	0.78	1.84	[1.36-2.33]	0	4.21	[1.36-2.33]	0	123 284	318 218	150 446
Pyr_As	IM	0.87	3.92	[3.13-4.74]	0	0.31	[0.21-0.40]	0	157 389	150 207	435 818
Ard_Cev	IM	0.64	2.14	[1.22-3.05]	299	2.61	[1.44-3.83]	351	32 274	97 936	110 861

Table sup3: Résultat DILS: tableau isolation avec migration