

DEVELOPPEMENT D'UN PANEL DE MARQUEURS SNP POUR DETECTER L'HYBRIDATION ENTRE LA CAILLE DES BLES ET LA CAILLE JAPONAISE.

Rouger Romuald¹, Brard-Fudulea Sophie¹, Thébault Noémie², Pitel Frédérique², Vignal Alain², Bed'hom Bertrand³, Rognon Xavier³, Lumineau Sophie⁴, Guémené Daniel¹

¹SYSAAF, INRA, Centre Val-de-Loire, 37380 Nouzilly, France ; ²GenPhySE, Université de Toulouse, INRA, ENVT, 31326 Castanet Tolosan; ³GABI, INRA, AgroParisTech, Université Paris-Saclay, Domaine de Vilvert, 78352 Jouy en Josas; ⁴Univ Rennes, Normandie Univ, CNRS EthoS UMR6552, 35042 Rennes.
romuald.rouger@inra.fr

RÉSUMÉ

La Caille des Blés (*Coturnix coturnix*) est une espèce migratrice couramment chassée en Europe. Des lâchers sont effectués dans les pays où cette pratique est autorisée afin de compléter les populations naturelles de cette espèce. De récentes études ont prouvé que certains individus de lâchers sont issus de l'hybridation avec la Caille japonaise (*Coturnix japonica*), espèce élevée pour sa chair et ses œufs. L'objectif de cet article est donc de proposer un jeu de marqueurs et une méthodologie statistique permettant de distinguer les individus *C. coturnix* des individus *C. japonica* et de leurs hybrides. A partir de données de séquençage issues de 20 *C. japonica* et 2 *C. coturnix*, 1.7 millions de SNP ont été détectés. Parmi ceux-ci, 192 marqueurs candidats ont été choisis comme potentiellement discriminants et testés sur des populations de référence. Comme attendu, la différenciation génétique estimée sur la base de ces marqueurs entre les deux espèces est forte ($F_{ST}=0.80$). Trente-deux marqueurs diagnostics de l'espèce ont pu être mis en évidence. Ils ont été inclus dans un panel de 96 marqueurs sélectionnés pour leur puissance statistique à discriminer les individus des deux espèces. Un traitement statistique *ad hoc* a été utilisé afin d'exploiter l'information produite par ce panel. Les résultats de simulations indiquent que la méthode permet de distinguer les individus issus de rétrocroisement jusqu'à la 3^{ème} génération avec un taux d'erreur d'assignation d'environ 1%. Testée sur des individus issus d'élevage, cette méthode démontre son utilité pour écarter les individus issus d'hybridation. Cette méthode est un nouvel outil permettant de contrôler le niveau d'introgression génétique de *C. japonica* dans les populations naturelles de *C. coturnix*.

ABSTRACT

The common quail (*Coturnix coturnix*) is a migratory species that is hunted in many countries across Europe. Restocking is performed in countries where it is allowed to supplement natural population of this species. Recent studies have shown that some of these restocking birds result from hybridization with the Japanese quail (*Coturnix japonica*) that is used for egg and meat production. The goal of this paper is to propose a set of markers and a statistical method allowing the discrimination of *C. coturnix*, *C. japonica* and their hybrids. Sequencing data obtained from 20 *C. japonica* and 2 *C. coturnix* allowed the detection of 1.7 millions SNP markers. From this list, 192 candidate markers were chosen as potentially discriminatory and tested on reference populations. As expected, the genetic differentiation between the two species estimated using these markers was high ($F_{ST}=0.80$) and thirty-two species-specific diagnostic markers were identified. These were included in a final panel of 96 markers selected for their statistical power to discriminate individuals from the two species. An *ad hoc* statistical method was used in order to use the information produced by this panel. Simulations showed that our method discriminates backcrossed individuals up to the third generation with an assignment error rate of approximately 1%. This method was also tested on farm animals where it showed its efficiency in removing hybrid individuals from the population. This method is a new promising tool to manage the level of genetic introgression of *C. japonica* in natural populations of *C. coturnix*.

INTRODUCTION

L'hybridation entre deux espèces proches est un phénomène couramment observé chez de nombreuses espèces animales dont les oiseaux (Randler 2002). Ses conséquences écologiques et évolutives peuvent être contrastées en ce qui concerne le maintien des espèces parentales. Ainsi l'hybridation peut être bénéfique en servant de pont entre deux espèces pour l'échange de gènes adaptatifs ou en contrant les effets de la dépression de consanguinité dans de petites populations. *A contrario*, la production importante d'hybrides peut impacter négativement la démographie des populations parentales si le succès reproducteur des hybrides est inférieur à celui des parents. Ce phénomène peut également aboutir au remplacement des populations parentales par l'hybride en cas d'hétérosis (i.e. vigueur hybride) (Todesco et al. 2016). Chez les oiseaux, une des causes communes d'hybridation est la mise en contact d'individus issus de populations naturelles avec des individus issus de populations domestiquées. Ainsi chez *Gallus gallus*, l'étude des plumages éclipses de spécimens de collection a mis en évidence que le type sauvage était progressivement remplacé dans son aire de répartition par des individus issus de l'hybridation avec le type domestiqué (Peterson and Brisbin 1999). Chez la perdrix, des hybrides entre la perdrix chukar (*Alectoris chukar*) et la perdrix rouge (*Alectoris rufa*) ont été utilisés afin de suppléer les effectifs de cette dernière, couramment chassée dans son aire de répartition. Néanmoins, la facilité des hybrides à se reproduire en milieu naturel couplé à la faible capacité des hybrides face au risque de prédation (Casas et al. 2011) sont susceptibles de mettre en danger certaines populations naturelles d'*A. rufa*. En réaction, un test ADN a été développé pour détecter les hybrides *Alectoris rufa x chukar*, test sur lequel se base maintenant la réglementation (Junta de Andalucía, 2016).

L'exemple de la Caille des Blés (*Coturnix coturnix*) est similaire. Cette espèce migratrice est largement chassée et des lâchers réguliers sont encore employés dans certains pays pour compenser les effectifs prélevés (e.g. Espagne, Italie ; Sanchez-Donoso et al. 2012). Afin d'améliorer les qualités zootechniques de ces individus de lâchers, l'hybridation avec la Caille japonaise (*Coturnix japonica*) est encore considérablement employée dans certains pays. La forte capacité des individus de lâchers à se reproduire par rapport aux individus sauvages se traduit par une introgression du génome de *C. japonica* à l'intérieur des populations sauvages de *C. coturnix* (Sanchez-Donoso et al. 2014). De plus la survie réduite des individus hybrides en environnement sauvage met en péril la survie des populations sauvages (Puigcerver et al. 2014). La détection des hybrides basée sur des critères phénotypiques (plumage, chant) est souvent difficile, notamment après plusieurs générations de

rétrocroisement. Afin de quantifier la présence d'individus hybrides en populations naturelles ou parmi les individus de lâchers, il est donc important de développer un test génétique *ad hoc*.

Le but de cette étude est donc de (1) proposer un panel de marqueurs moléculaires spécifiquement développés pour distinguer les individus *C. coturnix* des individus *C. japonica*; (2) adapter une méthode statistique permettant l'assignation d'individus dont le pedigree est inconnu.

1. MATERIELS ET METHODES

1.1. Séquençage et désignation des marqueurs SNP candidats.

Les séquences de 20 *C. japonica* issues de deux lignées sélectionnées de façon divergente sur leur comportement social (lignées S+ et S-, INRA Val-de-Loire) ainsi que celle de 2 *C. coturnix* (U. Rennes, CNRS) ont été obtenues à la suite du programme ANR SNP-BB (ANR 009-GENM-008) et mises à disposition pour ce projet. A partir de ces séquences, 1.7 millions de SNP ont pu être détectés. Une phase de nettoyage bio-informatique et d'analyses des fréquences alléliques a permis de réduire ce nombre à 41000 SNP potentiellement spécifiques à l'une ou l'autre des espèces, parmi lesquels 192 marqueurs candidats ont été choisis dans le but de couvrir le génome de *Coturnix sp.* de façon homogène.

1.2. Génotypage des populations de références et analyses préliminaires.

Quatre-vingt-six individus considérés comme appartenant à l'espèce *C. coturnix* et 91 individus considérés comme appartenant à l'espèce *C. japonica* ont été génotypés à chacun des 192 marqueurs candidats (Tableau 1). Soixante individus hybrides dont le niveau d'hybridation est connu et 24 individus provenant d'un élevage espagnol ont également été génotypés. Enfin, 3 individus provenant de l'étude de Chazara et al. 2010 et issus de prélèvements de chasse ont été inclus, parmi lesquels deux avaient été identifiés comme hybrides (Individus 3 et 4 de Chazara et al. 2010), le troisième ayant un profil génétique incertain (Individu 9 de Chazara et al. 2010) (Tableau 1).

De ce jeu de données ont tout d'abord été écartés les SNP présentant plus de 5% de données manquantes. Ensuite, les individus présentant plus de 5% de données manquantes ont été retirés. Cette phase de filtrage a abouti à l'obtention d'un jeu de données épuré comportant 171 marqueurs pour 232 individus (dont 66 *C. coturnix*, 81 *C. japonica*, 25 individus de statut incertain et 60 hybrides de niveau d'hybridation connu, Tableau 1).

Une Analyse en Composantes Principales (ACP) a été conduite sur les échantillons provenant des deux populations *C. coturnix* et *C. japonica* afin d'évaluer la confiance dans l'utilisation de ces derniers comme échantillons de référence. Les 96 meilleurs marqueurs

ont ensuite été sélectionnés sur la base de l'indice F_{st} de différenciation génétique (Nei 1987).

Tableau 1 Nombre et origine des échantillons utilisés dans la présente étude. n correspond au nombre d'individus génotypés, n_1 correspond au nombre d'individus ayant passé le filtre de qualité.

	n	n_1
<i>C. japonica</i>		
Sept lignées commerciales	71	61
Population INRA (Val-de-Loire, PEAT-Nouzilly)	10	10
Individus non domestiqués (Japon)	10	10
<i>C. coturnix</i>		
Prélèvement de chasse :		
France (Chazara et al. 2010)	27	27
Espagne	20	20
Maroc	10	5
U. Rennes 1, CNRS	29	14
<i>Statut incertain</i>		
Elevage espagnol	24	22
Individus issus de Chazara et al. 2010	3	3
<i>Hybrides de niveau d'hybridation connu obtenus à l'INRA Val de Loire (PEAT-Nouzilly)</i>		
Hybrides F1	20	20
Hybrides F2	20	20
Hybrides F7	20	20
Total	264	232

1.3. Méthode statistique

La méthode statistique utilisée ici est inspirée des méthodes classiques d'assignation individuelle à des classes d'hybrides (Montecinos et al. 2017; Campton and Utter 1985) et est une simplification de la méthode implémentée dans le logiciel NewHybrids (Anderson and Thompson 2002). Dans cette approche, nous considérons qu'un individu de pedigree inconnu peut être issu d'un nombre fini de classes d'individus (C_i) (Tableau 2).

Tableau 2 Dénomination des classes d'individus (C_i) utilisées dans cette étude. Le terme rétrocroisement se comprend ici comme le croisement d'un individu hybride avec un individu « Blé » ou « Jap ».

Dénomination	Description
« Blé »	Individus <i>C. coturnix</i>
« Jap »	Individus <i>C. japonica</i>
« F1 »	Croisement entre un individu « Blé » et un individu « Jap »
« Fx »	Croisement entre deux individus hybrides dont aucun ancêtre n'est issu de rétrocroisement
« BléB ₁ »	Rétrocroisement entre un individu « F1 » et un individu « Blé »
« BléB _i »	Rétrocroisement entre un individu « BléB _{i-1} » et un individu « Blé »
« JapB ₁ »	Rétrocroisement entre un individu « F1 » et un individu « Jap »

« JapB₁ » Croisement entre un individu « JapB_{i-1} » et un individu « Jap »

Soit un locus biallélique l dont les allèles sont A et a ; sachant les fréquences alléliques dans chacune des populations « Blé » et « Jap » de référence, il est possible d'exprimer la fréquence de l'allèle A (p_{C_i}) au sein de chaque classe :

$$p_{C_i} = 0,5 \times p_{C_j} + 0,5 \times p_{C_k} \quad (1)$$

Avec p_{C_j} et p_{C_k} les fréquences alléliques au sein des classes des deux parents j et k dont un individu de classe C_i est le produit. De même, les fréquences génotypiques attendues au sein de n'importe quelle classe sont :

$$p(AA)_{C_i} = p_{C_j} \times p_{C_k}$$

$$p(Aa)_{C_i} = p_{C_j} \times (1 - p_{C_k}) + (1 - p_{C_j}) \times p_{C_k} \quad (2)$$

$$p(aa)_{C_i} = (1 - p_{C_j}) \times (1 - p_{C_k})$$

Ainsi, pour un individu dont on connaît le génotype G_i à un locus il est possible de calculer la probabilité d'observer ce génotype dans une classe d'individu donnée C_i . On note cette probabilité $p(G_i|C_i)$. Sous hypothèse d'indépendance entre loci, la probabilité d'observer un génotype multilocus G à l'intérieur d'une classe donnée C_i est donc égale à :

$$p(G|C_i) = \prod_{l=1}^L p(G_l|C_i) \quad (3)$$

Où L correspond au nombre de loci auxquels chaque individu est génotypé.

Par utilisation de la formule de Bayes et en faisant l'hypothèse *a priori* qu'un individu a la même probabilité de provenir de chacune des classes d'individu, on peut déterminer la probabilité que l'individu de génotype G provienne de la classe C_i comme

$$p(C_i|G) = \frac{p(G|C_i)}{\sum_{m=1}^M p(G|C_m)} \quad (4)$$

Avec M le nombre de classes d'individus considérés. On se réfère à cette probabilité dans la suite de cet article comme étant la probabilité d'assignation du génotype G à la classe d'individus C_i . On considérera un individu comme appartenant à une classe si sa probabilité d'assignation à cette classe est supérieure à 0,90. L'implémentation de cette méthode a été effectuée à l'aide du logiciel R (R Core Team 2018).

1.4. Assignment d'individus hybrides produits *in silico*

Un jeu de données a été simulé afin de tester la performance du jeu de 96 marqueurs pour la détection d'individus hybrides. Ce jeu de données comporte 100 individus pour chacune des classes suivantes : Blé, Jap, F1, F2, BléB₁, BléB₂, BléB₃. Le génotype multilocus de chaque individu a été simulé en tenant compte des probabilités données en équation (1). La probabilité d'assignation de ces individus simulés a été calculée pour chacune des classes suivantes : « Blé », « Jap », « F1 », « Fx », « BléB₁ », « BléB₂ », « BléB₃ », « JapB₁ », « JapB₂ », « JapB₃ ».

1.5. Assignment d'individus réels hybrides ou de statut incertain.

La probabilité d'assignation des 60 individus hybrides de niveau connu et des 25 individus de statut incertain a été calculée en utilisant les mêmes classes que pour l'assignation des individus simulés.

2. RESULTATS ET DISCUSSION

2.1. Statistiques descriptives et sélection des 96 meilleurs marqueurs

Sur les 171 marqueurs retenus à la suite du génotypage, 32 marqueurs présentent un caractère diagnostique de l'espèce (i.e. l'un des allèles est fixé chez *C. coturnix* et l'autre allèle est fixé chez *C. japonica*). L'analyse en composantes principales basée sur l'ensemble des 171 marqueurs montre une nette distinction entre les deux populations de référence, sans individu de coordonnées intermédiaires (Figure 1). Les deux individus *C. coturnix* ségrégeant le long de l'axe 2 de l'ACP sont ceux dont le génome a été séquencé afin de définir les 192 marqueurs à tester, quelques marqueurs spécifiques à ces deux individus les différenciant des autres individus de l'espèce. Le F_{ST} multilocus entre *C. japonica* et *C. coturnix* calculé sur la base des 171 marqueurs est de 0,80. Les 96 meilleurs marqueurs ont été sélectionnés sur la base de leur F_{ST} (mini=0,69, maxi=1). A l'échelle de ces 96 marqueurs, le F_{ST} multilocus est de 0,93.

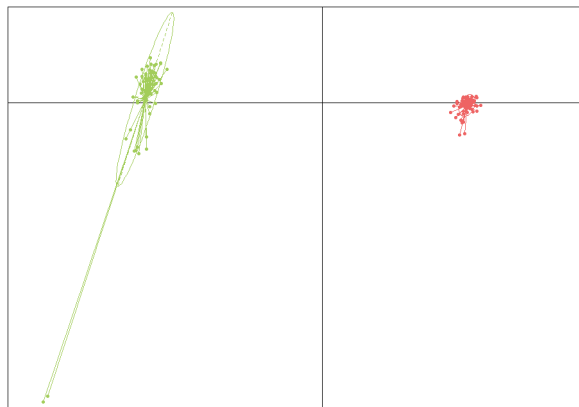


Figure 1 Analyse en composantes principales basée sur les génotypes des populations *C. coturnix* (en vert) et *C. japonica* (en rouge). Les axes des abscisses et des ordonnées expliquent respectivement 67,9% et 4% de l'inertie totale.

2.2. Assignment d'hybrides *in silico* aux différentes classes d'individus.

Parmi les 100 individus *C. coturnix* produits par simulation, tous sont correctement assignés à leur espèce avec des probabilités d'assignation supérieures à 0,90 (Figure 2). Les résultats sont similaires pour les individus *C. japonica*, F1 et F2. A quelques exceptions près (5 dans cette simulation), les individus BléB₁ sont également tous assignés à la bonne classe. Le nombre d'individus au statut ambigu augmente pour les

individus BléB₂ et BléB₃. Pour les individus BléB₂, une proportion non négligeable ont un statut ambigu entre la classe « BléB₃ » et « BléB₂ ». Le pattern est identique pour les individus BléB₃ dont la probabilité d'assignation à la classe « Blé » est quelquefois élevée. Néanmoins, cette probabilité ne dépasse le seuil d'assignation que dans de rares cas (Figure 2). La méthode présentée ici présente ses limites lors de la tentative d'assignation d'individus BléB₄ (données non présentées), les fréquences alléliques attendues dans la classe d'individu correspondante « BléB₄ » sont tellement proches des fréquences de la classe « Blé » que 96 marqueurs ne sont pas suffisant pour obtenir une distinction satisfaisante des individus des deux classes. En résumé, cette simulation démontre que la méthode présentée ici distingue les individus issus de rétrocroisement de 3^{ème} génération (« BléB₃ ») des individus *C. coturnix* avec un taux d'erreur d'assignation d'environ 1%.

2.3. Assignment d'individus réels hybrides ou de statut incertain.

Parmi les soixante individus hybrides de niveau connu, tous les individus ont été assignés sans ambiguïté à la classe dont ils sont censés être issus, à l'exception d'un seul (Figure 2). La haute probabilité d'assignation de cet individu F7 à la classe « JapB₁ » est probablement le fruit du hasard, les aléas de méiose dans sa généalogie ayant abouti à la mise à l'état homozygote de nombreux allèles *C. japonica* chez celui-ci.

Les deux individus considérés comme hybrides dans l'étude de Chazara *et al.* (2010) ont ici également été assignés à la classe hybride « Fx ». Le troisième individu au profil génétique incertain a lui été considéré comme appartenant sans ambiguïté à la classe « Blé » dans le présent article. Dans leur étude, Chazara *et al.* (2010) ont utilisé des marqueurs microsatellites qui n'étaient pas choisis sur la base de leur propriété discriminante entre *C. coturnix* et *C. japonica*. Le gain de précision obtenu grâce à l'utilisation du jeu de marqueurs spécifiques permet donc d'expliquer ce résultat.

Enfin, parmi les 22 échantillons provenant de l'élevage espagnol, seulement 13 individus sont assignés à la classe « Blé », 3 sont assignés à la classe « BléB₁ », 3 à la classe « BléB₂ », et 3 ont un statut ambigu entre la classe « BléB₁ » et la classe « BléB₂ ». Ce résultat démontre l'utilité de la méthode pour détecter la présence d'individus hybrides en populations naturelles et contrôler le statut génétique des individus de lâchers dans les pays où cette pratique est utilisée.

CONCLUSION

Notre étude met à disposition un outil moléculaire et statistique permettant de distinguer les individus *C. coturnix* de *C. japonica* et de leurs hybrides. Les études antérieures ayant eu le même objectif mettaient en œuvre des marqueurs non originellement développés dans ce but (Smith *et al.* 2018; Amaral *et al.* 2007;

Sanchez-Donoso et al. 2014). Cette étude est la première pour laquelle le choix des marqueurs a été explicitement effectué avec l'objectif de différencier les deux espèces.

Le lâcher de *Coturnix sp.* afin de compléter les populations naturelles est encore une pratique courante dans certains pays (Espagne, Italie). Jusqu'à présent aucun outil n'existait pour vérifier le statut génétique des individus utilisés lors de ces lâchers. Notre méthode permet ce contrôle et pourra être utile afin de limiter les risques d'introgessions génétiques de *C. japonica* vers les populations naturelles de *C. coturnix* (Smith et al. 2018).

De nombreuses pistes d'amélioration sont également envisagées pour raffiner la méthode (plus grand

nombre de marqueurs, description fine de la diversité génétique chez les populations naturelles de *C. coturnix*). En parallèle, une méthode d'estimation de l'introgession génétique au niveau populationnel et non plus individuel est en cours de développement. Ces deux méthodes complémentaires feront l'objet d'une future communication.

REMERCIEMENTS

Une partie du financement du projet a été supporté par AGENAVI (projet SNP-Cailles)

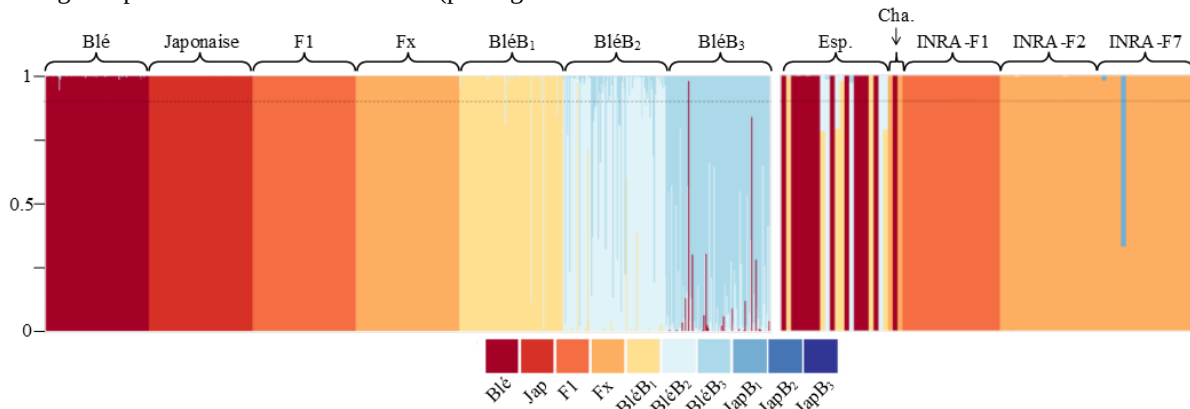


Figure 2 Graphique représentant les probabilités d'assignation aux différentes classes d'individus. Chaque barre verticale représente un individu. A gauche, les individus produits par simulation (de Blé à BléB₃) ; à droite les individus réellement génotypés (de Esp. à INRA-F7). (Esp. : Cailles issues d'un élevage espagnol, Cha. : Cailles considérées comme hybrides par Chazara *et al.* (2010)). La ligne horizontale en pointillé représente le seuil d'assignation placé à 0.9 dans cette étude.

REFERENCES BIBLIOGRAPHIQUES

- Amaral A., Grosso A.R., Chikhi L., Bastos-Silveira C., 2007. *Folia Zool.* 56(3):285-300
- Anderson E.C., Thompson E.A., 2002. *Genetics.* 160(3), 1217-29
- Campton D.E., Utter F.M., 1985. *Can. J. Fish. Aquat. Sci.* 42(1), 110-119
- Casas F., Mougeot F., Sánchez-Barbudo I., Dávila J. A. Viñuela, J., 2011. *Biol. Invasions.* 14(2), 295-305
- Chazara O., Minvielle F., Roux D., Bed'hom B., Feve K., Coville J.-L., Kayang B.B., Lumineau S., Vignal A., Boutin J.M., Rognon X., 2010. *Conserv. Genet.* 11(3), 1051-1062
- Montecinos A.E., Guillemain M.-L., Couceiro L., Peters A.F., Stoeckel S., Valero M., 2017. *Mol. Ecol.* 26(13), 3497-3512
- Nei M., 1987. In : *Molecular Evolutionary genetics.* Columbia University Press, pp512
- Peterson A.T., Brisbin I.L., 1999. *Bird Conserv. Int.* 9,387-394
- Puigcerver M., Sanchez-Donoso I., Vilà C., Sardà-Palomera F., García-Galea E., Rodríguez-Teijeiro J. D., 2014. *Biol. Conserv.* 171,74-81
- R Core Team, 2018. <https://www.R-project.org/>.
- Randler C., 2002. *Anim. Behav.* 63(1), 103-119.
- Sanchez-Donoso I., Huisman J., Echegaray J., Puigcerver M., Rodriguez-Teijeiro J. D., Haile F., Vila C., 2014. *Front. Ecol. Evol.* 2,15.
- Sanchez-Donoso I., Vilà C., Puigcerver M., Butkauskas D., Caballero de la Calle J. R., Morales-Rodríguez P. A., Rodríguez-Teijeiro J. D. 2012. *PLoS ONE.* 7(6)
- Smith S., Fusani L., Boglarka B., Sanchez-Donoso I., Marasco V., 2018. *Eur. J. Wildlife Res.* 64(5):51.
- Todesco M., Pascual M.A., Owens G.L., Ostevik K.L., Moyers, B.T., Hübner S., Heredia S.M., Hahn M.A., Caseys C., Bock D.G., Rieseberg, L. H., 2016. *Evol. Appl.* 9(7), 892-908.
- Junta de Andalucía, 2016. https://www.juntadeandalucia.es/export/drupaljda/Copia%20de%20Borrador%20de%20Orden%20de%20certificaci%C3%B3n%20gen%C3%A9tica%20Corregido-2_1.pdf