



UNIL | Université de Lausanne

Unicentre

CH-1015 Lausanne

<http://serval.unil.ch>

Year : 2024

Using genomics to measure the extent and consequences of inbreeding

Lavanchy Eléonore

Lavanchy Eléonore, 2024, Using genomics to measure the extent and consequences of inbreeding

Originally published at : Thesis, University of Lausanne

Posted at the University of Lausanne Open Archive <http://serval.unil.ch>

Document URN : urn:nbn:ch:serval-BIB_7447169ECAA87

Droits d'auteur

L'Université de Lausanne attire expressément l'attention des utilisateurs sur le fait que tous les documents publiés dans l'Archive SERVAL sont protégés par le droit d'auteur, conformément à la loi fédérale sur le droit d'auteur et les droits voisins (LDA). A ce titre, il est indispensable d'obtenir le consentement préalable de l'auteur et/ou de l'éditeur avant toute utilisation d'une oeuvre ou d'une partie d'une oeuvre ne relevant pas d'une utilisation à des fins personnelles au sens de la LDA (art. 19, al. 1 lettre a). A défaut, tout contrevenant s'expose aux sanctions prévues par cette loi. Nous déclinons toute responsabilité en la matière.

Copyright

The University of Lausanne expressly draws the attention of users to the fact that all documents published in the SERVAL Archive are protected by copyright in accordance with federal law on copyright and similar rights (LDA). Accordingly it is indispensable to obtain prior consent from the author and/or publisher before any use of a work or part of a work for purposes other than personal use within the meaning of LDA (art. 19, para. 1 letter a). Failure to do so will expose offenders to the sanctions laid down by this law. We accept no liability in this respect.



UNIL | Université de Lausanne

Faculté de biologie
et de médecine

Département d'Ecologie et Evolution

Using genomics to measure the extent and consequences of inbreeding

Thèse de doctorat ès sciences de la vie (PhD)

présentée à la

Faculté de biologie et de médecine
de l'Université de Lausanne

par

Eléonore LAVANCHY

Biologiste diplômée de l'Université de Lausanne

Jury

Prof. Sven Bergmann, Président
Prof. Jérôme Goudet, Directeur de thèse
Prof. Roman Arguello, Expert
Prof. Daniel Wegmann, Expert

Lausanne
(2024)



UNIL | Université de Lausanne

Faculté de biologie
et de médecine

Ecole Doctorale

Doctorat ès sciences de la vie

Imprimatur

Vu le rapport présenté par le jury d'examen, composé de

Président·e	Monsieur	Prof.	Marc	Robinson-Rechavi
Directeur·trice de thèse	Monsieur	Prof.	Jérôme	Goudet
Expert·e·s	Monsieur	Prof.	Roman	Arguello
	Monsieur	Prof.	Daniel	Wegmann

le Conseil de Faculté autorise l'impression de la thèse de

Eléonore Lavanchy

Master - Maîtrise universitaire ès Sciences en comportement, évolution et conservation,
Université de Lausanne

intitulée

**Using genomics to measure the extent
and consequences of inbreeding**

Lausanne, le 12 avril 2024

pour le Doyen
de la Faculté de biologie et de médecine

Prof. Marc Robinson-Rechavi

Table of contents

Table of contents	2
Remerciements	4
Summary	6
Résumé	7
General Introduction	9
Thesis Outline	18
Chapter I: Effect of reduced genomic representation on using runs of homozygosity for inbreeding characterization	20
Chapter II: Detecting inbreeding depression in structured populations	37
Chapter III: Too big to purge: no evidence of purging in the islands of the European barn owl (<i>Tyto alba</i>)	66
Chapter IV: Inbreeding depression in the Swiss barn owl (<i>Tyto alba</i>) population	94
General Discussion.....	117
References (General Introduction and Discussion).....	123
Supplementary Material Chapter I	129
Supplementary Material Chapter II.....	154
Supplementary Material Chapter III	195
Supplementary Material Chapter IV	209
Annex I.....	217
Annex II.....	265

Annex III269

Annex IV281

Remerciements

J'aimerais en premier lieu remercier mon superviseur, Jérôme Goudet. Merci de m'avoir supervisée avec patience, merci de m'avoir fait confiance en me laissant de la liberté scientifique. Mais le plus grand merci concerne ton humanité: merci de m'avoir écoutée, encouragée et soutenue dans les moments difficiles. Si j'ai réussi à arriver jusqu'ici c'est en grande partie grâce à toi, Merci.

Je remercie ensuite chaleureusement mon comité de thèse : merci au président Sven Bergmann ainsi qu'aux experts Roman Arguello et Daniel Wegmann pour le temps consacré à la lecture de ma thèse.

Un énorme merci aux membres du groupe Goudet que je ne remercierais jamais assez. Au cours de ces 5 années ils ont été comme une deuxième famille qui m'a toujours aidée dans les périodes difficiles, qui m'a accompagnée et aidé à grandir en tant que scientifique mais aussi en tant qu'être humain. Merci à mes deux mentors originels qui m'ont accueillie en tant que toute nouvelle PhD : Ana Machado et Tristan Cumer. Merci à Ana qui a été mon exemple : qui m'a accueillie au sein du groupe à mon arrivée, qui m'a appris les bases la gestion d'un PhD et m'a soutenue lors de mes premières années de thèse. Merci à Tristan qui m'a accompagnée en tant que collègue tout au long de ces cinq années, qui m'a (presque) tout appris, qui a relu l'intégralité des textes que j'écrivais lors de la première moitié de mon doctorat et qui a toujours été disponible pour discuter et m'aider lorsque j'avais le moindre problème malgré son emploi du temps plus que chargé. Merci à Alexandros Topaloudis, pour son humour incroyable, son soutien quotidien et les discussions scientifiques intéressantes. Merci à Isabela do O avec qui j'ai immédiatement accroché et dont la gentillesse est difficilement égalable. Merci de m'avoir écoutée et soutenue dans les périodes difficiles, d'avoir pris sur toi des charges de travail lorsque tu savais que j'étais stressée, d'avoir relu beaucoup de mes textes et d'avoir toujours été disponible pour tenter de régler mes problèmes scientifiques. Merci à Clara Castex pour ton inébranlable jovialité, pour m'avoir toujours écoutée et soutenue, sans jugement, quel que soit le moment où j'arrive dans ton bureau ! Merci à Anna Hewett qui est arrivée sur ma fin de thèse et qui a largement contribué à l'avancée rapide de mon dernier chapitre. Merci d'avoir collaboré avec moi sur ce dernier chapitre, de m'avoir expliqué plusieurs concepts scientifiques et de m'avoir proposé ton aide pour soulager ma charge de travail lors de cette fin de thèse. Merci à Hugo Corval pour ton humanité et ton calme qui sont agréables dans notre bureau ! Merci

d'avoir toujours été disponible pour discuter quel que soit le sujet. Merci à Angelica Pulido pour m'avoir écoutée parler des difficultés et du stress de fin de thèse et pour m'avoir encouragée et remonté le moral.

Merci au reste de mes amis de l'UNIL qui tout comme mon groupe ont largement égayé ces 5 années et m'ont apporté un soutien inconditionnel : ma « chosen family » : Julie Guenat, Giulia Perroud et Sagane Dind qui m'accompagnent depuis le master, avant même le début de ma thèse, merci aussi à Aijuan Liao, Miya Pan, Kai-Hsiu Chen, Thu Nguyen, Laurie Ançay, Estelle Millet, Kim Schalcher, Anne-Lyse Ducrest, Molly Baur, Sarah Schmidt, Camille-Sophie Cozzarolo, Iris Prigent ainsi que celles et ceux que j'ai probablement oublié (désolée) : vous êtes des personnes incroyables. Merci aussi à mes amis en dehors de l'UNIL, pour vos encouragements et votre amitié: Fanny Tang, Julia Landrein, Line Martin et Sandra Bonvin.

Merci à ma famille : Noémie Lavanchy, Christiane Meister et Laurent Lavanchy de m'avoir soutenue dans mon choix d'étude mais aussi de m'avoir rassurée et constamment rappelé que si c'était trop dur pour moi je pouvais arrêter à tout moment et que c'était OK. Merci de m'avoir écoutée me plaindre et de m'avoir fait à manger à pleins d'occasions (c'est important).

Finalement, je voudrais remercier particulièrement Christian de Guttry pour m'avoir accompagnée tout du long de cette aventure : lors du choix de continuer sur une thèse, d'avoir discuté de longs moments avec moi de quel sujet serait intéressant, d'avoir relu l'intégralité de mes textes scientifiques lors de ma première année. Merci de m'avoir toujours soutenue inconditionnellement, de m'avoir remonté le moral dès que je me sentais désemparée et d'avoir littéralement assuré le 99% des tâches lors des derniers mois de ma thèse. Sans toi, cette aventure aurait n'aurait pas été pareil. Merci

Summary

Inbreeding can have negative effects on individuals, a phenomenon called inbreeding depression. It can occur as a result of mating between close relatives such as siblings and cousins but also as a result of mating between more distant relatives. The latter is often observed in populations with few individuals. Due to human activities, both the number of species as well as the number of individuals within populations is declining. The ability to quantify the inbreeding of individuals and its effects is therefore crucial for conserving biodiversity. In this thesis, we first evaluated methods for inbreeding and inbreeding depression quantification. We then applied these methods to study the inbreeding status of the European barn owl (*Tyto alba*) using genome-wide sequencing data.

Sequencing the complete genome of many individuals can be expensive and many popular and cheaper methods exist for sequencing small fractions of the genome. First, we compared two methods for quantifying individual inbreeding based on genomic data. We demonstrated that although the popular method can be used when the genome has been sequenced at high density, it produces biased estimates at low densities. On the contrary, a less popular and more computationally demanding method gave accurate results with genomic data at both high and low densities. We then proposed a new statistical method for quantifying inbreeding depression strength in data with a strong genomic structure. We suggested taking into account the genetic relatedness among all the individuals and showed that it allows meaningful estimation of inbreeding depression when the standard method results in highly biased estimations.

We then studied genome-wide sequencing data from 502 barn owls from all around Europe. We demonstrated that populations living on islands are more inbred and enriched in deleterious variants mostly because of smaller effective population sizes resulting in lower selection efficiency. Finally, we combined genomic and phenotypic data from 3,085 barn owls from Switzerland and we quantified the negative effect of inbreeding on bill length in barn owls juveniles.

This thesis highlights that inbreeding quantification must be done using accurate data and methods. It becomes even more apparent when it comes to humans' health or the conservation of endangered species (which are particularly susceptible to inbreeding).

Résumé

La consanguinité est le résultat de croisements entre individus apparentés. Malgré le fait que ses effets délétères sur les individus sont documentés depuis longtemps, elle reste, chez l'humain, une pratique courante dans certaines cultures. C'est aussi un phénomène commun chez d'autres espèces : avec l'augmentation constante des pressions anthropogéniques sur les populations sauvages ces dernières années— telle la dégradation des habitats, par exemple –, la taille de nombreuses populations diminue drastiquement. Cela conduit à une augmentation de la parenté entre les individus, et donc à une augmentation de la consanguinité moyenne des populations.

Il est important de pouvoir correctement quantifier le degré de consanguinité ainsi que ses effets dans une population. Dans la première partie de cette thèse, nous évaluons différentes méthodes pour estimer le degré de consanguinité des individus, ainsi que pour mesurer ses conséquences. La plupart des méthodes existantes sont basées sur des marqueurs moléculaires. Séquencer l'intégralité du génome d'un individu coûte cher, et il existe plusieurs procédés permettant de n'en séquencer qu'une partie. Nous comparons la capacité à détecter de la consanguinité entre une méthode facile à utiliser et très populaire et une autre méthode moins utilisée et plus compliquée à mettre en place. Nous montrons que les deux méthodes fonctionnent très bien lorsque l'intégralité du génome est séquencée. En revanche, lorsque la fraction de génome séquencée est faible, seuls les résultats obtenus avec la seconde méthode sont pertinents. Nous proposons ensuite une nouvelle approche statistique pour estimer l'effet de la consanguinité sur des données contenant une forte structure génétique. Nous proposons de prendre en compte la parenté entre tous les individus. Nous comparons différentes méthodes pour calculer cette parenté, ainsi que différentes méthodes pour estimer le degré de consanguinité des individus. Nous montrons que notre approche permet d'estimer correctement l'effet de la consanguinité dans un jeu de données contenant de la structure génétique, ce qui n'est pas le cas avec la méthode classique.

Dans la seconde partie de cette thèse, nous utilisons des données génomiques d'effraies des clochers (*Tyto alba*) afin de quantifier la consanguinité de différentes populations européennes. Nous montrons que les populations des îles sont plus consanguines que celles vivant sur le continent, et que cette consanguinité est principalement due à la petite taille de leurs populations et à leur plus forte isolation. Finalement, nous quantifions l'effet négatif de la consanguinité

dans la population suisse et montrons qu'elle est liée à une réduction de la taille du bec chez les jeunes effraies.

En résumé, cette thèse démontre l'importance d'utiliser de bonnes données et méthodes pour estimer la consanguinité ainsi que ses effets négatifs. Cela est particulièrement important lorsque les enjeux concernent la santé des êtres humains ou l'évaluation du risque d'extinction de populations ou d'espèces menacées.

General introduction

Diversity in Nature

Nature is characterized by an extreme diversity. The International Union for Conservation of Nature (IUCN) reported 2,130,023 species in December 2021, consisting of 73,883 vertebrates, 1,491,386 invertebrates, 423,373 plants, and 141,381 fungi and protists. While these numbers may seem large, they represent less than 5% of the current estimates of the number of living species on Earth [65, 94]. The estimated species number exceeds 8.75 million, ranging from “simple” photosynthetic unicellular organisms to extremely complex multicellular organisms such as giant squids found in the ocean, succulent plants surviving in deserts or us humans. In addition to interspecific diversity, there is also considerable diversity within species, although it is not always entirely apparent. Both genetics and environment can contribute to an individual’s observable physical characteristics, known as phenotype, including (but not limited to) appearance and behavior. For instance, the capacity of humans to taste the bitterness of the phenylthiocarbamide chemical compound is so strongly genetically determined that it was used as a sort of paternity test during the middle of the 20th century [7]. On the contrary, the southern subspecies of the gaudy commodore butterfly (*Precis octavia sesamus*) found in West Africa exists in two completely different morphs: a blue winter morph and an orange summer morph (figure 1). The temperature at which the larva develops entirely determines the adult morph [32].

The gaudy commodore winter morph
Precis octavia sesamus



The gaudy commodore summer morph
Precis octavia sesamus



Figure 1: The seasonal morphs of the gaudy commodore butterfly (*Precis octavia sesamus*). The left panel represents the blue winter morph while the right panel shows the orange summer morph. Both images were taken from Wikipedia and are credited to Purves M. and Svdmolten respectively.

However, the constant anthropogenic pressures such as habitat reduction, increased hunting pressure and rapid climatic changes, lead to significant reduction in both inter- and intraspecific diversity and more generally in many population sizes [12, 13, 11, 71]. Decreases in population size can lead to population crash and eventually extinction. Indeed, population size is linked to population fitness via the interplay of drift strength and selection efficiency. In large populations, selection overcomes drift: it maintains deleterious alleles at low frequencies and increases the frequency of beneficial mutations. When the population size is reduced, drift can overcome the effect of selection if the strength of selection is lower than $1/2N_e$. As a result, deleterious alleles may rise in frequency [44] or even reach fixation [28]. In addition to lowering the overall population fitness by increasing the frequency of deleterious alleles, a small population size also leads to higher relatedness among individuals. This, in turn, increases the probability of mating between related individuals.

Inbreeding

The term inbreeding refers to the mating of related individuals. It can occur when these individuals are closely related (i.e., the time of coalescence between both parents is short), such as when siblings or cousins mate, which will be hereafter referred to as consanguineous mating. The process of breeding using closely related individuals is extremely prevalent in domestic species, where breeders select animals with similar phenotypes (hence often closely related) for the purpose of producing offspring with the preferred phenotypes [77, 52]. The practice of consanguineous mating is also prevalent in humans and has even been encouraged in some periods and regions, including the Habsburg family in Europe between 1450 and 1750 [9] or in north and sub-Saharan Africa, the Middle East, and west, central, and south Asia today [3]. Inbreeding may also occur in small isolated populations as a result of more ancient relatedness (i.e. the time of coalescence between both parents is long). In this regard, island populations are perfect examples: they are often colonized in few events and their gene flow with other populations is restricted. Consequently, most of their members are descended from the same ancestors. Such levels of more ancient inbreeding have been observed in the black-footed rock wallaby of Barrow Island in Western Australia [26], the island populations of platypus off the coast of South Australia [29], the barn owl in Crete and Cyprus in Europe [57] and the hihi passerine in New-Zealand [24]. In fact, since all organisms descended from a common ancestor, all individuals are ultimately related. As a result, all measures of inbreeding describe an individual's inbreeding status with respect to a base population. This base population can be

the last known ancestor [95], a putative founder population [96] or the current population itself [93].

Regardless of whether inbreeding occurs through recent or ancient coalescence events, its effect on individuals is the same: it increases the proportion of genome which is autozygous in inbred individuals [16]. Yet, the coalescence time is relevant because the more ancient it is, the smaller the fraction of the genome that is autozygous [91]. Here, I would like to emphasize that there is a difference between being homozygous and being autozygous. A homozygous individual carries two copies of the same allele at a locus regardless of which ancestral allele they come from. If both copies of the allele actually come from the same ancestral allele, then the individual is both homozygous and autozygous at this locus and both alleles are identical-by-descent (IBD). On the contrary, if an individual carries two copies of the same allele but these copies do not come from the same ancestral allele, then the individual is homozygous at this locus but not autozygous and the two alleles are said to be identical-by-state (IBS). This difference is illustrated in figure 2.

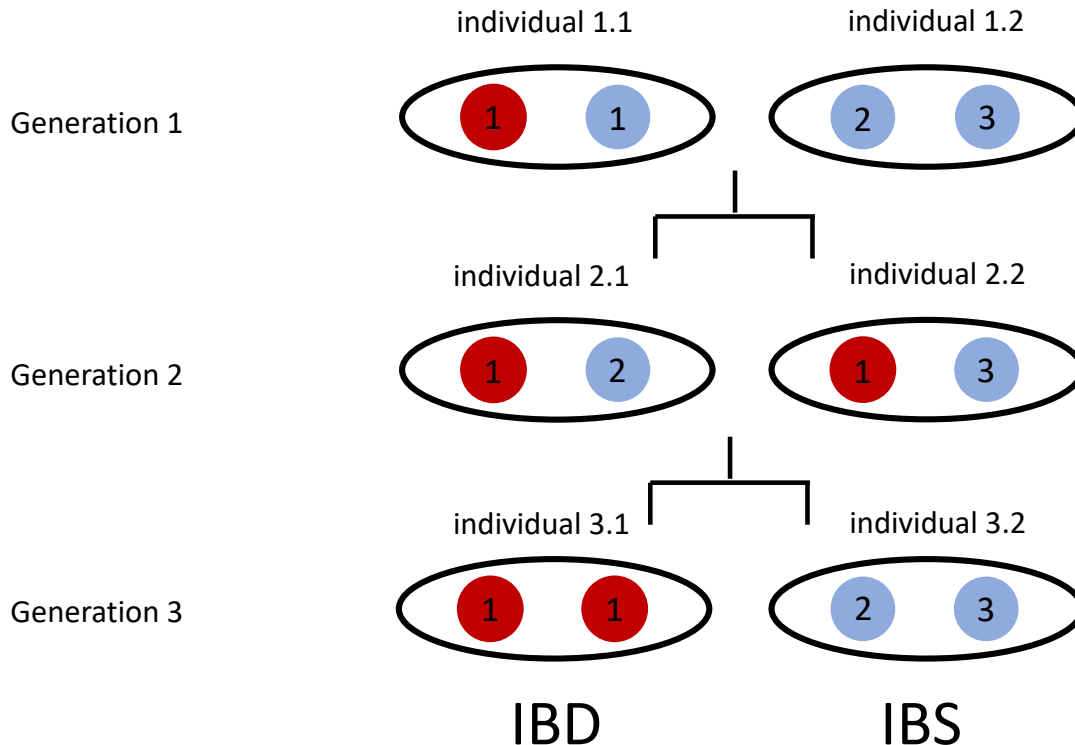


Figure 2: The difference between identical-by-descent (IBD) and identical-by-state (IBS). The figure depicts a biallelic locus in which a blue and a red allele are segregating. In the first generation, we have

two individuals: individual 1.1 is heterozygous and carries one copy of the red allele (1) and one copy of the blue allele (1). Individual 1.2 is homozygous and carries two copies of the blue allele (2 and 3). The two individuals mate and produce two offspring (2.1 and 2.2) and both are heterozygous. The first parent (1.1) transmitted the same red allele (1) to both of them. On the contrary, the second parent (1.2) transmitted a different blue allele to both offspring (individual 2.1 received the blue allele (2) while individual 2.2 inherited the blue allele (3)). These two siblings mate and create two new offspring (individuals 3.1 and 3.2). The first siblings (3.1) received the red allele (1) from both parents. Since both of these red alleles came from the same ancestral red allele (1) (from individual 1.1), individual 3.1 is both homozygous and autozygous at this locus and its two red alleles are IBD. Individual 3.2 received two blue alleles (2 and 3). However, these blue alleles do not originate from the same ancestral allele (despite coming from the same ancestor 1.2). Consequently, individual 3.2 is homozygous at this locus but not autozygous and its alleles are IBS. Of course, this is only because our base population are individuals 1.1 and 1.2 and the two blue alleles in individual 3.2 would eventually be IBD if we go back in time long enough.

Inbreeding depression and purging

Increase in inbreeding has been linked to fitness reduction in many taxa [15, 87, 62, 43, 40, 44, 68, 86, 24, 35] - a phenomenon called inbreeding depression and which was first documented by Darwin in plants [18]. Inbreeding depression is stronger for tightly linked fitness-related traits such as survival in sheep [86], killer whales [43] and hihi [24] or life-time breeding success in red deers [40], hihi [24] and helmeted honeyeater [35]. However, inbreeding depression has also been observed in morphological traits such as human height [62] and house sparrow mass and bill length [68]. The major mechanism responsible for inbreeding depression was identified in 2009 by Charlesworth and Willis [14]. Even though selection actively acts against deleterious alleles, recessive deleterious alleles are difficult to eliminate since they are only expressed in homozygous individuals. Consequently, these alleles often segregate at low frequencies in populations [73]. As mentioned above, inbreeding increases autozygosity (and hence homozygosity) which will lead to the expression of these recessive deleterious alleles in inbred individuals, thereby reducing their fitness. Traditionally, inbreeding depression is measured as the difference in fitness (or any trait) between inbred and outbred individuals: the slope of a (generalized linear) regression of the focal trait on the individuals' inbreeding coefficients.

Paradoxically, even though inbreeding can lead to fitness reduction, long-term inbreeding and small population size can also lead to the elimination of recessive deleterious alleles - a phenomenon called purging [31, 30, 25, 78]. As mentioned above, inbred individuals have reduced fitness, hence reducing their chance of transmitting their alleles to the next generation. If the majority of the population is inbred, most deleterious alleles segregating within the population will be expressed and hence eliminated by selection. Since selection efficiency is directly related to allele deleteriousness, purging usually occurs for highly deleterious alleles

[31]. Therefore, even populations that have undergone purging may still be enriched in mildly and lowly deleterious alleles and susceptible to inbreeding depression [31, 30, 25, 47].

How to measure inbreeding

Due to the devastating effects inbreeding can have on individuals, quantifying it and its consequences is crucial. Inbreeding is often measured at the individual level through inbreeding coefficients - denoted F . The first method for quantifying inbreeding was proposed by Sewall Wright in 1922. He suggested counting the number of loops in pedigrees [95] (the pedigree-based inbreeding coefficient is called F_{PED}). F_{PED} measures the probability that two alleles within an individual are IBD [58] (in respect to the pedigree founders). While this approach is still valid today, it presents two major disadvantages: i) it requires an active record of all mating events, which is difficult to achieve in practice, particularly in wild populations, and ii) what F_{PED} really measures is the expected probability that an allele is IBD. Indeed, during meiosis, because of Mendelian segregation and recombination stochasticity, the genomes of individuals are not composed of exactly 25% of each grandparent's genome and there is randomness in which part of these grandparents' genomes they will receive. The "true" (i.e. genomic-based) average coancestry between full-siblings is 0.25 (as estimated with the pedigree) but its 95% confidence interval spans between 0.204 and 0.296 [33, 38, 85]. The difference between the expected fraction of genome that is IBD and the fraction that is actually IBD is illustrated in figure 3.

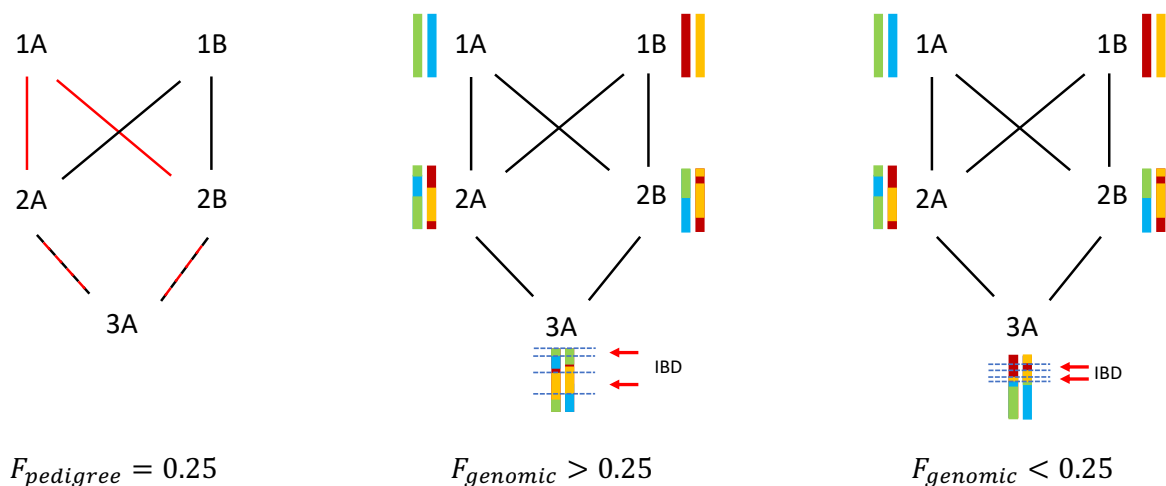


Figure 3: The difference between pedigree-based and genomic-based inbreeding coefficients. This figure represents the mating between full siblings (2A and 2B) and the results in their corresponding offspring (3A). In the left part of the chart, the family genealogy is shown, for counting loops: there are two possible loops, topping with the two grandparents (1A and 1B). On the right side of the figure, two different offspring

possibilities can be seen based on different allele segregation and recombination events. Chromosomes are represented by colored rectangles. The chromosomes of individuals 2A, 2B, and 3A are mosaics of the chromosomes of their ancestors (1A and 1B).

With the advancement in sequencing technologies, it is now possible to sequence individual genomes and, therefore, measure the “true” genomic inbreeding coefficient of individuals. Numerous studies showed that genomic measures of inbreeding are better than pedigree-based for the reasons mentioned above [41, 45, 1]. However, many different approaches have been proposed to measure genomic-based inbreeding coefficients: i) relying on the correlation between uniting gametes [96], the mean allele-sharing of individuals [93] or the comparison between observed and expected heterozygosity [75]. While correlations between different measures are acceptable in homogeneous populations, they are low (and even negative) in populations with strong genetic structure [99]. The underlying reason is that the inbreeding estimates mentioned above (except for the allele-sharing-based estimate) rely on allelic frequencies, and the allelic frequencies of subpopulations may differ greatly from those of the metapopulation. In addition, all these measures actually treat each position of the genome independently, but we know that DNA is transmitted from parents to offspring via large chromosomal chunks. Consequently, McQuillan and collaborators proposed in 2008 to use large homozygous segments within individuals’ genomes - called runs of homozygosity (ROHs) - as a proxy for IBD segments [63].

Runs of homozygosity

There is evidence that ROHs are good proxies for detecting IBD segments within an individual [51, 39]. By quantifying the fraction of genome within ROHs, one can estimate a ROHs-based inbreeding coefficient (F_{ROH}) [63]. In addition, their distributions (i.e. lengths and numbers) can be used to infer populations’ history and past demography [63, 46, 69, 10]. Indeed, the length of two IBD segments (and hence a ROH) is directly linked to the time of coalescence between both segments [91, 85]. The presence of long ROHs in a population indicates consanguineous mating. Conversely, a population with a large number of small ROHs indicates ancient relatedness among individuals, which may reflect a small effective population size or past bottlenecks. This is illustrated in figure 4 (from Ceballos *et al.* 2018, *Nature reviews genetics* [10]).

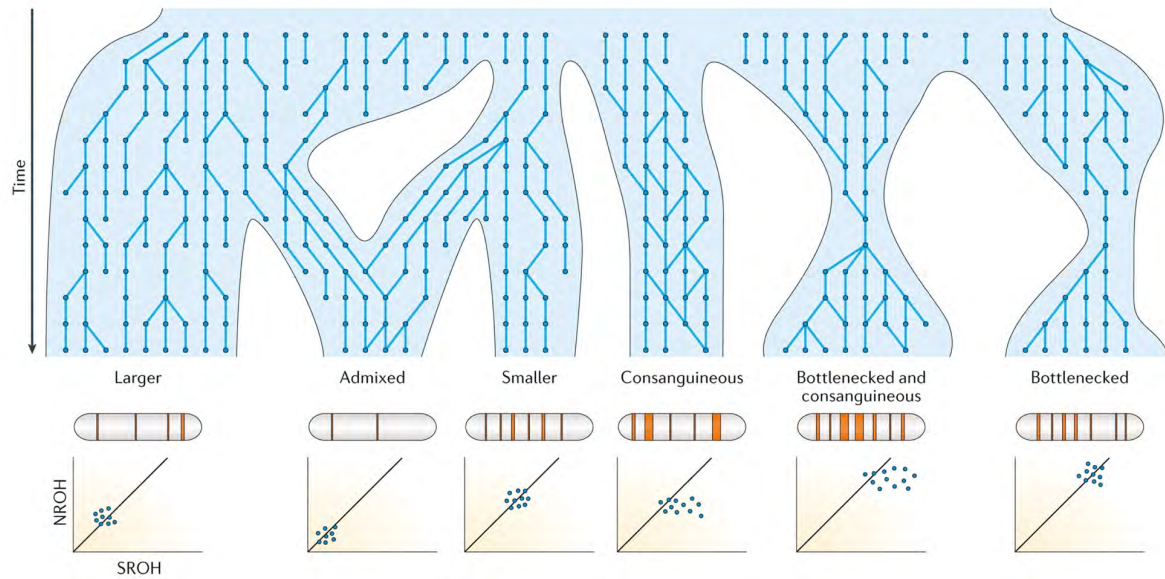


Figure 4: The demographic history of six diverse hypothetical populations is represented in the upper part of the plot. Figure (and legend) are from Ceballos *et al.* 2018, *Nature reviews genetics* [10]. Representative pedigrees are indicated by dark blue lines connecting individuals (dots), loops show inbreeding and the population size is represented by the width of the light blue areas. Thus, bottlenecks are shown by a narrowing, which necessarily reduces the number of ancestral lineages that are present in the population; conversely, larger populations contain more ancestral lineages. Admixture is shown by a confluence of two hitherto separate populations and mating between the pedigree lineages therein. The consequences of each demographic scenario are illustrated below: schematic chromosomes showing the typical distribution of runs of homozygosity (ROH) in each and at the bottom a plot of the sum total length of ROH (SROH) versus the total number of ROH (NROH) expected in each scenario. As can be seen, the burden of ROH relates to the size of the population, with smaller populations having more and longer ROH than larger populations. Admixture brings together different haplotypes and typically reduces the number of ROH to very few short ROH, whereas bottlenecks increase the number of ROH, which are typically still relatively short. Consanguinity, on the other hand, adds a small number of very long ROH for those who are the offspring of cousin marriage, thus also increasing the variance in the sum of ROH, visible as a right shift in the NROH versus SROH plot. Some populations are both bottlenecked and practice consanguineous marriage, hence having many short and some long ROH, resulting in the highest burden of ROH.

On the importance of simulated data in biology

Quantifying specific parameters (such as inbreeding levels here) in order to better understand populations and their histories is the core of applied biology. To develop, improve, and test methods for parameter estimation, it is necessary to know what the true value of the parameter we are trying to quantify is. With empirical data, this is obviously not possible. Therefore, researchers have used simulated data (e.g. genomic or phenotypic) for which they set the true value of the parameters in order to test methods. Importantly, generating meaningful simulated data requires a thorough understanding of the underlying biological mechanisms (such as the

transmission of genetic material through generations in our case). However, as thorough our understanding is, simulations will never perfectly mimic what happens in nature. Consequently, it is also important to validate methods with empirical data.

The barn owl

The barn owl (*Tyto sp.*) is a nocturnal raptor found on all continents except Antarctica [48, 20, 92]. The species consists of several subspecies that are distributed throughout the continents. The common barn owl clade (*Tyto alba*) is widespread in Africa and throughout Europe (including its islands) [57, 56, 17, 92] and there is low genetic structure within the European mainland populations [56, 17]. This wide distribution with various population sizes makes it ideal to investigate and compare inbreeding levels of individuals. Barn owls live in open habitats (such as fields) [84] and feed mostly on small rodents [80]. During the breeding season (which lasts from March to October), barn owls produce large broods (up to a dozen eggs), primarily fed by the male as the female warms the eggs and chicks until they are old enough to keep themselves warm [81, 88, 67]. Afterwards, the female assists with feeding. For more than 30 breeding seasons, a portion of the Swiss barn owl population has been the subject of a long-term study. In the study area, mating events are recorded. Several morphological traits, such as color, mass, wing, tarsus, and bill length are also measured and blood samples are collected from each bird caught in the study area (illustrated in figure 5).



Figure 5: The barn owl (*Tyto alba*). Top panel: An adult barn owl returns to its nest box with its prey. Bottom left panel: The wing length of an adult barn owl is being measured in the field. Bottom right panel: The weight of a juvenile barn owl is being measured in the field.

Thesis outline

In this thesis, I combine simulated and empirical data to study the quantification and consequences of inbreeding and inbreeding depression. In the first part, we compare software and statistical methods for quantifying both inbreeding and inbreeding depression. In the second half of this thesis, we used empirical barn owl (*Tyto alba*) data to estimate individuals' inbreeding status and inbreeding depression strength.

Throughout chapter one, we compare two approaches for identifying runs of homozygosity (ROHs): a fast and widely used observational-based approach and a more computationally challenging and less widely used model-based approach. We simulate whole-genome sequencing (WGS) data and examine the accuracy of both methods in identifying IBD segments when various fractions of the genome are available. The results suggest that both methods perform accurately with high-density genomic data. Although observational approaches are commonly used, they tend to overestimate the number of large ROHs and underestimate the number of small ROHs when small portions of the genome are available. Model-based approaches, however, perform well even with small fractions of the genome.

In the second chapter, we propose and test a novel method for quantifying inbreeding depression in populations with a strong genetic structure. We propose using the genetic relatedness matrix (which summarizes the kinship between individuals) to account for the non-independence of such data in (general) linear models. We compare different relatedness matrices and inbreeding coefficients and show that the most commonly used matrix performs poorly under these conditions. We demonstrate, however, that another matrix (based on allele-sharing) can be used to correct for structure and provide accurate estimates of inbreeding depression strength.

In the third chapter, we examine the inbreeding status of 502 barn owls from various European populations. More specifically, we compare mainland populations with islands populations both in terms of inbreeding levels and enrichment in deleterious alleles. We demonstrate that island populations are more inbred than mainland populations, and this inbreeding is primarily caused by ancient relatedness due to small effective sizes. In addition, we show that islands' populations are enriched in deleterious alleles, indicating a lower selection efficiency and a stronger effect of drift.

Lastly, in chapter four of this thesis, we quantify the effect of inbreeding depression on various traits in the Swiss barn owl population. We find evidence for moderate inbreeding depression in juveniles, suggesting that they are more vulnerable to inbreeding depression than adults.

Chapter I: Effect of reduced genomic representation on using runs of homozygosity for inbreeding characterization

Authors: Eléonore Lavanchy^{1,2} and Jérôme Goudet^{1,2}

¹ Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland

² Swiss Institute of Bioinformatics, University of Lausanne, Lausanne, Switzerland

Chapter status: published on the 10th of January 2023 in *Molecular Ecology Resources*:

<https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.13755>

Authors contribution: EL and JG conceptualized the study. EL conducted the study, performed all simulations and analyses and drafted the manuscript. EL and JG wrote and revised the manuscript.

RESOURCE ARTICLE

Effect of reduced genomic representation on using runs of homozygosity for inbreeding characterization

Eléonore Lavanchy^{1,2}  | Jérôme Goudet^{1,2} ¹Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland²Swiss Institute of Bioinformatics, University of Lausanne, Lausanne, Switzerland**Correspondence**Eléonore Lavanchy and Jérôme Goudet, Department of Ecology and Evolution, Biophore, University of Lausanne, CH-1050 Lausanne, Switzerland.
Emails: eleonore.lavanchy@unil.ch; jerome.goudet@unil.ch**Funding information**

Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: 31003A_179358

Handling Editor: Kimberly J. Gilbert**Abstract**

Genomic measures of inbreeding based on identical-by-descent (IBD) segments are increasingly used to measure inbreeding and mostly estimated on SNP arrays and whole-genome sequencing (WGS) data. However, some softwares recurrently used for their estimation assume that genomic positions which have not been genotyped are nonvariant. This might be true for WGS data, but not for reduced genomic representations and can lead to spurious IBD segments estimation. In this project, we simulated the outputs of WGS, two SNP arrays of different sizes and RAD-sequencing for three populations with different sizes and histories. We compare the results of IBD segments estimation with two softwares: runs of homozygosity (ROHs) estimated with PLINK and homozygous-by-descent (HBD) segments estimated with RZooRoH. We demonstrate that to obtain meaningful estimates of inbreeding, RZooRoH requires a SNPs density 11 times smaller compared to PLINK: ranks of inbreeding coefficients were conserved among individuals above 22 SNPs/Mb for PLINK and 2 SNPs/Mb for RZooRoH. We also show that in populations with simple demographic histories, distribution of ROHs and HBD segments are correctly estimated with both SNP arrays and WGS. PLINK correctly estimated distribution of ROHs with SNP densities above 22 SNPs/Mb, while RZooRoH correctly estimated distribution of HBD segments with SNPs densities above 11 SNPs/Mb. However, in a population with a more complex demographic history, RZooRoH resulted in better distribution of IBD segments estimation compared to PLINK even with WGS data. Consequently, we advise researchers to use either methods relying on excess homozygosity averaged across SNPs or model-based HBD segments calling methods for inbreeding estimations.

KEYWORDS

homozygous-by-descent, identical-by-descent, inbreeding, reduced genomic representations, runs of homozygosity

1 | INTRODUCTION

Inbreeding is defined as mating between relatives and has been observed across many taxa including humans (Bittles &

Black, 2010; Ceballos et al., 2018), livestock (Forutan, Ansari Mahyari, et al., 2018; Kim et al., 2013; Peripolli et al., 2017, 2018), wild animal populations (Åkesson et al., 2016; Huisman et al., 2016; Kardos et al., 2018; Keller & Waller, 2002) and

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2023 The Authors. *Molecular Ecology Resources* published by John Wiley & Sons Ltd.

plants (Kariyat & Stephenson, 2019; Keller & Waller, 2002; Zhang et al., 2019). Its quantification and the understanding of its deleterious consequences – called inbreeding depression – are central in many areas of biology, from human genetics to conservation biology (Keller & Waller, 2002). Indeed, increase in genome autozygosity has been associated with diseases, such as schizophrenia (Keller et al., 2012; Lencz et al., 2007) and Alzheimer's disease (Ghani et al., 2015; Nalls et al., 2009) as well as fitness costs in animals (Åkesson et al., 2016; Huisman et al., 2016) and plants (Menges, 1991; Zhang et al., 2019).

Individual levels of inbreeding are quantified with inbreeding coefficients (F). Traditionally, inbreeding was measured by counting the size and number of loops in pedigrees (F_{PED}) (Wright, 1922) a method with several downsides: (i) it estimates the expected individual coefficient which can differ from the realized individual coefficient due to recombination stochasticity and Mendelian segregation (Carothers et al., 2006; Franklin, 1977; Hill & Weir, 2011); (ii) it assumes founders of the pedigree are unrelated and noninbred; (iii) pedigrees must be correctly recorded which is extremely difficult in wild populations, although genetic data might be used to (re)construct links (Huisman, 2017; Jones & Wang, 2010). With the advancements in high throughput sequencing technologies it became possible to estimate with sufficient accuracy genomic-based F , and several studies have shown molecular estimates to be more accurate than pedigree-based estimates (Alemu et al., 2020; Kardos et al., 2015; Keller et al., 2011; Wang, 2016). Many different genomic-based F have been proposed, such as F_{HOM} (Chang et al., 2015), F_{AS} (Weir & Goudet, 2017), F_{UNI} and F_{GRM} (both described in Yang et al., 2011) but there is still no consensus on which is the most accurate (Alemu et al., 2020; Caballero et al., 2020; Goudet et al., 2018; Nietlisbach et al., 2019; Yengo et al., 2017). These estimates quantify average excess single nucleotide polymorphism (SNP) homozygosity or correlation between uniting gametes and treat all SNPs independently. However, parents transmit DNA to their offspring in large chromosomal segments rather than each base independently. Consequently, it has been suggested that measures of inbreeding should be based on identical-by-descent (IBD) segments rather than individual SNPs (McQuillan et al., 2008). Hence, a new F was proposed by McQuillan et al., (2008). This coefficient intends to quantify the proportion of IBD segments in the genome. From this point onward, we will call the true fraction of genome within IBD segments: F_{IBD} and its estimations (i) F_{ROH} when estimated from runs of homozygosity (ROHs) with observational-based approaches and (ii) F_{HBD} when estimated from homozygous-by-descent (HBD) segments from model-based approaches.

McQuillan et al. (2008) proposed to use ROHs, long consecutive homozygous segments, as a proxy for these IBD segments. ROHs were first described by (Broman & Weber, 1999) and shown to be ubiquitous in humans (Ceballos et al., 2018; Gibson et al., 2006; Pemberton et al., 2012) and across many different taxa (Kardos et al., 2018; Liu et al., 2020; Saremi et al., 2019). F_{ROH} is calculated as the proportion of the genome within ROHs and

several studies demonstrated that it was a reliable estimator of inbreeding (Alemu et al., 2020; Caballero et al., 2020; Nietlisbach et al., 2019). In addition to quantifying inbreeding, distribution of IBD segments (i.e., lengths and numbers) can inform about a population's past demography and history (Ceballos et al., 2018; Kirin et al., 2010; Pemberton et al., 2012): long segments reflect recent coalescence events while smaller segments indicate more distant coalescence and, if in high proportion, a history of small effective population size. Finally, IBD segments can be used for identifying rare deleterious recessive variants responsible for deleterious phenotypes by homozygosity mapping, which in short compares islands of IBD segments between affected and unaffected individuals (Alkuraya, 2013; Hildebrandt et al., 2009; Stoffel et al., 2021; Wang et al., 2009).

Two different methods for IBD segments detection are recurrently used in the literature: observation and model-based approaches (Ceballos et al., 2018). The most common method is a fast observation-based method (Ceballos et al., 2018) implemented in PLINK (Chang et al., 2015; Purcell et al., 2007). It makes use of a sliding window to identify continuous homozygous stretches, with a minimum size defined by the user, used as proxy for IBD segments. The other family of methods is model-based, and has been implemented in RZooRoH (Bertrand et al., 2019; Druet & Gautier, 2017, 2022), BEAGLE (Browning & Browning, 2010) and BCFTools (Narasimhan et al., 2016). It relies on hidden Markov models (HMM) and directly infers HBD segments from the genotypes by considering the distance between two markers, the mutation rate and even the recombination map if available. Consequently, these methods do not require a minimum threshold on segment length. HMM methods are computationally demanding (Ceballos et al., 2018) and a previous study suggested that PLINK outperformed HMM methods both in terms computation time and ROHs detection accuracy with simulated whole-genome-sequencing (WGS) data (Howrigan et al., 2011). However, few HMM methods were available at that time and the authors did not investigate the robustness of PLINK to genotyping errors. Observation-based approaches were designed for WGS data and assume that the region between two SNPs are entirely homozygous. However, many studies performing ROHs analyses with PLINK used reduced genomic representation techniques: often SNP arrays (Bjelland et al., 2013; Bosse et al., 2012; de Jong et al., 2020; Forutan, Mahyari, et al., 2018) where specific SNPs, chosen based on their position, effect on phenotype or minor allele frequency (MAF), are targeted and genotyped. ROHs have also been called with restriction-site associated DNA sequencing (RAD-sequencing) data, by cutting the genome near enzymes cutting sites and selecting and sequencing fragments based on their size. With both SNP arrays and RAD sequencing, only a fraction of the genome is sequenced resulting in a partial representation of the total polymorphism. Since PLINK assumes that genomic positions not included in the SNPs set are nonvariant, we expect that it will falsely consider nonsequenced heterozygous loci as homozygous which can lead to spurious ROHs detection. On the contrary,

the HMM approach from Leutenegger et al. (2003), which models the genome as a mosaic of HBD and non-HBD segments and from which most current model-based approaches follow, was initially developed for SNP arrays. These models do not treat nonsequenced genomic regions as homozygous but as missing data. However, model-based approaches are rarely used for detecting IBD segments with reduced genomic data (but see Alemu et al., 2020; Dutsch et al., 2021; Sole et al., 2017). In addition, no precise benchmarking with large sample size has been performed on comparing how the different methods behave with these reduced genomic data compared to WGS data and precise guidelines such as which method is suitable with which data are missing.

In addition to the fraction of genome captured, we hypothesize that the effective size and level of polymorphism in a population might also affect the capacity of the different methods to accurately detect IBD segments. Small and inbred populations tend to harbour higher numbers of long such segments easier to accurately detect with reduced representations as the missing positions are more likely to be homozygous. Larger populations will tend to harbour many small IBD segments (Ceballos et al., 2018; Kirin et al., 2010) harder to detect when only a fraction of the polymorphism is available since these small segments require lots of nearby SNPs to reach the minimum density threshold for an accurate detection (Kardos et al., 2015; Sole et al., 2017; Zhang et al., 2015). On the other hand, larger populations will harbour higher levels of polymorphism and thus higher numbers of SNPs resulting in an increased SNPs density for the same fraction of genome sequenced with RAD-sequencing.

Here, we use simulated data to compare the performance of PLINK and RZooRoH with two reduced genomic representations – SNP arrays and RAD-sequencing – and WGS. We compare both softwares output to the true IBD segments extracted from the simulated data. We hypothesize that the quality of detection depends on SNP density. In addition, since model-based approaches take into account the distance between each SNP, we predict that they will perform better when dealing with sparse data (Druet & Gautier, 2017). We show that both detection methods can be used to correctly estimate IBD segments with SNP arrays and RAD-sequencing providing that a sufficient proportion of the genome has been sequenced. This proportion varied between IBD segments detection methods and population sizes: the model-based method implemented in RZooRoH as well as the large population require a substantially smaller fraction of the genome to obtain correct inbreeding and distribution of IBD segments estimates.

2 | MATERIALS AND METHODS

All scripts used in this project are available on GitHub: <https://github.com/EleonoreLavanchy/ROHsReducedRep>. A general workflow of the study can be found in Figure 1 and additional details about the simulations and analyses performed can be found in the Supporting Information.

2.1 | Simulations

We simulated two hermaphroditic populations ($N = 1000$ and $N = 10,000$) using SLiM3, a forward-in-time individual-based simulation software (Haller & Messer, 2019). We used a “non-Wright-Fisher” model with nonfixed population sizes and overlapping generations (Haller & Messer, 2019). Population size was regulated via a patch carrying capacity where individuals were removed based on their overall fitness at the end of each simulation cycle. Individuals' fitness decreased with age which varied between 0 and 3: older individuals had higher probabilities to die. Individuals were able to reproduce from the age of 1 and selfing was not allowed. For each individual, its mate was chosen among the other individuals based on their age (with older individuals less likely to be chosen) and on their pedigree-based coancestry with the focal individual (related individuals had higher chances to be chosen). This resulted in a population mostly practicing random mating but ensured that some inbreeding would occur at each generation. We simulated 10 replicates for both population sizes and each simulation lasted for 1000 reproductive cycles. We used a human-like genetic map with a nonhomogenous recombination rate simulated with FREGENE as described in Chadeau-Hyam et al. (2008). This resulted in genomes of 3000 centimorgans (cM). Individuals from both populations carried 30 chromosomes each 100Mb long. The burnin were performed via recapitation in msprime (Kelleher et al., 2016). All mutations were added at the end of the simulation (after the burnin) based on a human-like mutation rate of 2.5×10^{-8} per site per generation (Nachman & Crowell, 2000).

At the end of the entire simulation process, we performed a random stratified sampling to ensure that the individuals used in subsequent analyses would cover the entire range of inbreeding. We are aware that this scheme is rare and hard to apply empirically but it allowed us to investigate whether the entire spectra of inbreeding was correctly estimated. Whenever possible, we subsampled 20 individuals with F_{PED} between 0 and 0.1, 20 individuals with F_{PED} between 0.1 and 0.2, 20 individuals with F_{PED} between 0.2 and 0.3, 20 individuals with F_{PED} between 0.3 and 0.4 and 20 individuals with F_{PED} between 0.4 and 0.5. The average (\pm SD) number of sampled individuals per replicate were 87.3 ± 5.03 for the small population and 67.40 ± 4.09 for the large population. The lower number of individuals subsampled in the large populations are because they contained fewer individuals with high F . At the very end, the mean (\pm SD) number of SNPs per simulated population was $1.6 \times 10^6 \pm 2.0 \times 10^4$ SNPs for the small populations and $1.6 \times 10^7 \pm 1.7 \times 10^5$ SNPs for the large populations.

2.2 | SNPs subsampling

In order to investigate the effect of reduced genomic representations on identical-by-descent (IBD) segments estimation, we mimicked different sequencing techniques by subsampling SNPs from whole-genome data. We simulated both RAD-sequencing and two SNP arrays of different sizes.

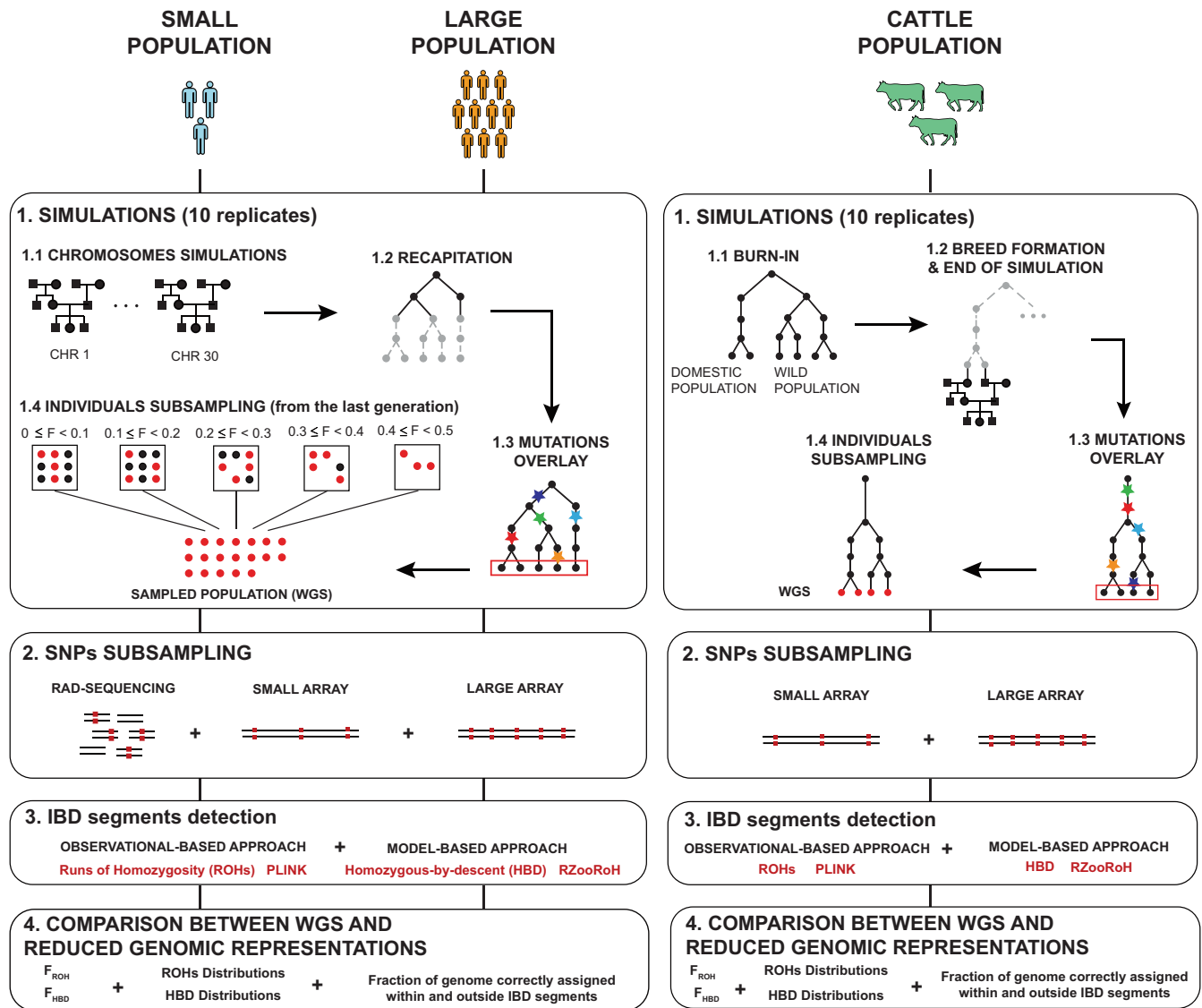


FIGURE 1 General workflow of the study. Simulations were first performed in SLiM3. Within one population, each chromosome shared the same pedigree. Burnin, recapitation and mutation overlay were then performed in msprime. Single nucleotide polymorphism (SNP) subsampling was performed with bedtools for RAD-sequencing and homemade python script for both arrays. Runs of homozygosity (ROHs) were called with PLINK and homozygous-by-descent (HBD) segments with RZooRoH. For distribution of ROHs and HBD segments, segments were divided into six length classes. The true identical-by-descent (IBD) segments were extracted from simulated TreeSequences. Fractions of genomes correctly and incorrectly assigned within and outside IBD segments were then estimated as the overlap between ROHs or HBD segments and these true IBD segments

RAD-sequencing uses restriction enzymes to digest the genome in small fragments, which are then selected on size and sequenced (Andrews et al., 2016). Consequently, these fragments are not homogeneously distributed along the genome. For this purpose, we randomly selected 500 base-pair (bp) fragments (Andrews et al., 2016) using bedtools version 2.29 (Quinlan, 2014). Afterwards, SNPs within these windows were subsampled using --bed function from VCFTools (Danecek et al., 2011). Given that the proportion of the genome sequenced with RAD-sequencing varies greatly depending on the organism and on the restriction enzymes used, we varied this number of fragments so that they covered between 0.05% and 10% of the genome, and between 0.002% and 1%, for the small and the

large populations respectively. We did so because these percentages resulted in similar SNPs densities between the small and large populations and because SNPs density is an accurate indicator for the accuracy of runs of homozygosity (ROHs) or homozygous-by-descent (HBD) segments detection (see results). In addition, we subsampled different percentages of genome (resulting in different SNP densities) between PLINK and RZooRoH. This is because RZooRoH requires smaller SNP densities compared to PLINK to reach the same accuracy of IBD segments estimation. We performed 100 replicates for each RAD-sequencing subsampling percentage.

To simulate SNP array sequencing, we mimicked two arrays initially developed for cattle and widely used for ROHs analyses: the

Illumina BovineSNP 50 beadchip (~50,000 SNPs) hereafter "small array" - and the Illumina BovineHD BeadChip - (~777,000 SNPs) hereafter "large array". Common features of both arrays are the homogenous distances between SNPs and the focus on common SNPs, hence, we first filtered our WGS data on MAF 5%. We then selected windows with size corresponding to the median distances between SNPs in the real arrays - 40kb for the small array and 3 kb for the large array - and selected the SNP with higher MAF within each window (if at least one SNP was present). We use the term "small array" for what is usually considered as a medium-density array in the literature.

2.3 | IBD segments estimation: ROHs And HBD segments

We compared two methods for IBD segments detection to investigate whether we observe a difference in their capacity to handle reduced genomic data. We chose one observational-based approach - the --homozyg method, implemented in PLINK (Chang et al., 2015) - and one model-based approach - the RZooRoH method, implemented as a R package (Bertrand et al., 2019; Druet & Gautier, 2017, 2022). PLINK makes use of a sliding window to identify homozygous segments, used as a proxy for IBD segments. Consequently, PLINK results will be called ROHs. On the contrary, RZooRoH models autozygous segments directly from the genotypes and its results will be referred to as HBD segments.

For ROHs detection with PLINK, we varied parameters according to the SNP density in the reduced data set as proposed by Kardos et al. (2015) and performed by Duntsch et al. (2021). In particular, we varied the window size (--homozyg-window-snp) as well as the minimum SNP density (--homozyg-density) and number of SNPs (--homozyg-snp) required for a homozygous segment to be called a ROH. We required lower numbers of SNPs for low SNPs densities data sets. We also varied the maximum number of heterozygous SNPs allowed in a ROH (--homozyg-het). On the contrary, we fixed some parameters: we authorized 1 heterozygous SNPs per window (--homozyg-window-het = 1), maximum 1 Mb in between two adjacent SNPs (--homozyg-gap = 1000) and a minimum ROH size of 100 kb (--homozyg-kb = 100). These parameters were consistent for every replicate per subsampling method. A more precise description and justification of how each parameter value was chosen according to the SNP density can be found in the Supporting Information (pages 3 to 5; Table S1 and Table S2; Figure S8 and Figure S9).

We called HBD segments with the RZooRoH package version 0.3.1 with a four HBD classes model with rates (R) equals 10, 100, 1000 and 10,000 for the HBD classes and 10,000 for the non-HBD class. These HBD classes correspond to different coalescence event ages: the rate corresponds to the expected number of generations since the coalescence event divided by 2 (i.e., 5, 50, 500 and 5000 generations ago, respectively). For each of these classes, the expected length of the HBD segments are defined as $1/R$ [in M]: 10 cM, 1 cM, 0.1 cM and 0.01 cM. Even though we chose a model with few

HBD classes, we expect that these classes cover all IBD segments length as the variances associated to these average lengths are extremely large (Speed & Balding, 2015). We used a value of 5×10^{-5} for genotype uncertainty, which represents the probability that any allele mutated in one of the ancestors in the last 1000 reproductive events: $\#meiosis \times mutation\ rate = 1000 \times 2 \times 2.5 \times 10^{-8}$.

2.4 | SNPs-independent measures of inbreeding: F_{HOM}

To test the performance of a SNPs-independent based F , we estimated F_{HOM} , implemented in the --het method from PLINK for all SNPs densities presented for RAD-sequencing in this manuscript and for WGS. F_{HOM} , F_{ROH} and F_{HBD} have different definitions and different assumptions about the 'base-population'. F_{HOM} aims at identifying excess homozygosity relative to a random-mating population. F_{ROH} aims at identifying IBD segments relative to an ancient "base-population" and the minimum length threshold chosen will set how far this base-population is (in coalescence time) compared to our current population. Finally, F_{HBD} does not use a minimum size threshold but identifies IBD SNPs by including the allelic frequencies into the emission probabilities of the HMM.

2.5 | True IBD segments

We compared the estimations obtained from both softwares and all fractions of genome subsampled to the true fraction of genome within IBD segments. We choose to consider a segment IBD if both haplotypes coalesced less than 100 reproductive cycles ago (independently of their length). Since we have overlapping generations and four age classes in our model, these 100 time-steps correspond to 25 generations. In supplementary material, we also benchmark PLINK and RZooRoH results with IBD segments coalescing less than 1000 time-steps (i.e., 250 generations) ago to account for more ancient coalescence events. This was done with the tmrca method from the tskit module in python.

2.6 | Statistical analysis

For PLINK, we estimated individual ROHs-based inbreeding coefficient (F_{ROH}) as the proportion of the genome within ROHs:

$$F_{ROH} = \frac{\sum Length_{ROH}}{genome\ length}$$
 (McQuillan et al., 2008). For RZooRoH, we estimated F_{HBD} as the average of posterior HBD probabilities across all markers in the data set (the @realized) as suggested in the RZooRoH documentation. We then compared these to the true fraction of genome within IBD segments (F_{IBD}).

Since distribution of IBD segments can inform about the population history (Ceballos et al., 2018), we divided these segments into six length classes following Kirin et al. (2010): (i) between 100kb and

2 Mb (i.e., between 0.1 cM and 2 cM), (ii) between 2 Mb and 4 Mb (i.e., between 2 cM and 4 cM), (iii) between 4 Mb and 6 Mb (i.e., between 4 cM and 6 cM), (iv) between 6 Mb and 10 Mb (i.e., between 6 cM and 10 cM), (v) between 10 Mb and 16 Mb (i.e., between 10 cM and 16 cM) and (vi) larger than 16 Mb (i.e., larger than 16 cM). Distribution of IBD segments are represented as the mean total length per individual among simulation and subsampling replicates. For the sake of comparison and because we can benchmark these with the true distributions of IBD segments, we used these length classes with both softwares even if they are traditionally used with PLINK as RZooRoH models and partitions HBD segments according to the rates chosen when constructing the model. Consequently, with RZooRoH we used the “most probable” distribution of HBD segments detected with the Viterbi algorithm rather than the average SNP probabilities of belonging to any length class. We then compared these distributions to the distribution of true IBD segments extracted from the tree sequences.

We use four metrics to evaluate the accuracy of ROHs detection for each subsampling technique: (i) the fraction of genome correctly assigned within IBD segments (true-IBD), that is, ROHs or HBD segments which were detected and which were IBD; (ii) the fraction of genome correctly assigned outside IBD segments (true-non-IBD), that is, genomic regions which were not classified as ROHs nor HBD segments with neither softwares and were not IBD segments; (iii) the fraction of genome inappropriately assigned within IBD segments (false-IBD), that is, genomic regions which were classified as ROHs or HBD segments but were not IBD; (iv) the fraction of genome inappropriately assigned outside IBD segments (false-non-IBD), that is, IBD segments which were not assigned as ROHs nor HBD segments. We compared ROHs and HBD segment estimation to the true IBD segments for every individual in every replicate, subsampling method and simulation. We then averaged individual's fractions among simulation and subsampling replicates to obtain one measure per subsampling event.

2.7 | Additional simulations

We performed an additional batch of simulations based on a real, 57 years deep, cattle pedigree from Walloon beef cattle. We used a genetic map estimated from male Holstein cattle by Qanbari and Wittenburg (2020). In the simulation, a domestic population ($N_e = 1500$) got separated from a large wild population ($N_e = 50,000$) 10,000 generations ago with a migration rate of 3×10^{-5} (Frantz et al., 2020). To mimic the strong selective pressure which occurred during breed formation 200 generations ago and which resulted in high levels of inbreeding (Frantz et al., 2020), 200 individuals were randomly selected from the domestic population and used as founders for the rest of the simulation. The remaining 200 generations were then simulated in SLiM3 from these 200 founders. As the real pedigree was only covering the last 57 years, a first round of simulations was run to obtain 200 generations-deep simulated pedigree, which was then used to complete the real pedigree by assigning a genealogy from the simulated pedigree to each founder from the

real pedigree. At the end, only the individuals from the real pedigree were kept for the analyses.

Since we showed with the first batch of simulations that the accuracy of ROHs detection with RAD-sequencing depends on the proportion of genome subsampled, we only mimicked SNP arrays-like subsampling for these simulations. We did so as previously described. At the end of the sliding windows process, we obtained a lower number of SNPs than expected as some windows did not contain any SNPs. Additional SNPs were chosen randomly (but still with $MAF > 0.05$) to account for empty windows and to reach the same number of SNPs as in real arrays.

Concerning IBD segments detection, we used the same parameters as we did for the small and large populations with PLINK and both SNP arrays. However, with WGS data we increased the maximum number of heterozygous SNPs authorized in a ROH (--homozyg-het) to 64 to optimize distribution of ROHs estimation. Concerning RZooRoH, we used the exact same model as before. For this cattle-like population, we compare the estimations obtained from both softwares to IBD segments coalescing less than 1000 (rather than 100) time-steps ago (independently of their length). This is because in this population, a large part of the inbreeding comes from old coalescence events: we show in Figure S10 that there is poor concordance between both softwares estimation and IBD segments coalescing less than 100 reproductive cycles ago.

2.8 | SNP density

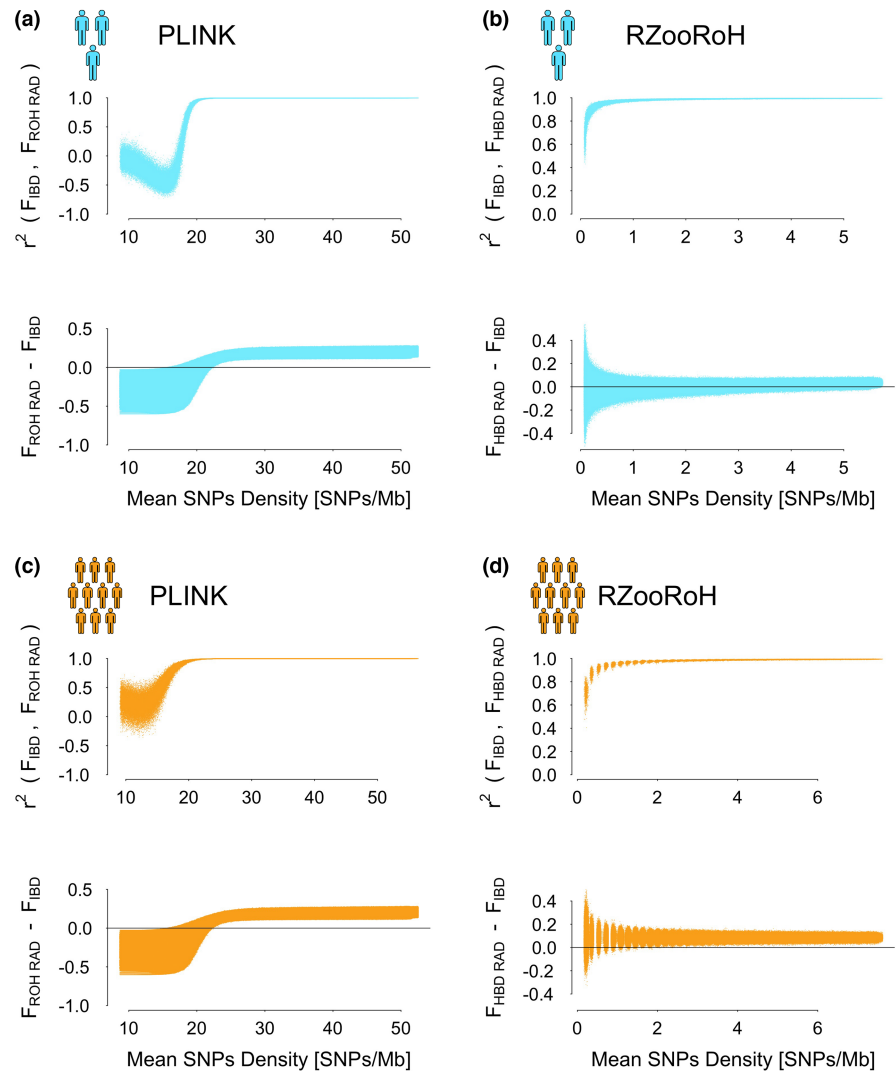
We calculated SNP densities with VCFtools (Danecek et al., 2011) method: --SNPdensity as the number of SNPs per each windows of 1 Mb. We then estimated the mean SNP density of each replicate as the mean density among the windows.

3 | RESULTS

3.1 | F_{ROH} and F_{HBD}

We used simulated genomes to investigate the influence of different sequencing techniques, IBD calling methods and population size on IBD segments detection. Figure 2 shows the correlation (r^2) between F_{IBD} (the true fraction of genome within IBD segments coalescing less than 100 reproductive cycles ago and calculated from simulated data) and its estimations: F_{ROH} (estimated with PLINK, Figure 2a,c) or F_{HBD} (estimated with RZooRoH, Figure 2b,d) according to SNP density as well as the absolute difference between the estimated $F_{ROH \text{ or } HBD}$ and F_{IBD} according to the same SNP density. Figure 2 shows F_{ROH} and F_{HBD} can be correctly estimated with reduced genomic representations providing that a sufficient fraction of the genetic variation is captured. To retain conserved ranking among individual F (represented by a correlation of one between $F_{ROH \text{ or } HBD}$ and F_{IBD}), PLINK (Figure 2a,c; Table S3) required a SNP density of 22 SNPs/Mb, 11 times higher compared to

FIGURE 2 Correlation (r^2) and difference between F_{IBD} and F_{ROH} (panels a and c) or F_{HBD} (b and d) estimated with RAD-sequencing data according to single nucleotide polymorphism (SNP) density in the reduced data set for both the small (a and b) and large (c and d) populations. Pearson's correlations were estimated per simulation and subsampling replicate. The difference was calculated by subtracting F_{IBD} from F_{ROH} or F_{HBD} per simulation and subsampling replicate. (a) Small population; runs of homozygosity (ROH) have been called with PLINK. (b) Small population; homozygous-by-descent (HBD) segments have been called with RZooRoH. (c) Large population; ROHs have been called with PLINK. (d) Large population; HBD segments have been called with RZooRoH.



RZooRoH (Figure 2b,d; Table S3) which only required two SNPs/Mb. Interestingly, these minimum SNP densities were similar between the two population sizes suggesting that SNP density might be a key metric for assessing the accuracy of F_{IBD} . With a SNP density below 20 SNPs/Mb, PLINK resulted in negative correlations between F_{ROH} and F_{IBD} in the small population (Figure 2a). On the other hand, correlations were always higher than 0.5 with RZooRoH (Figure 2b,d). We should stress that the conservation of inbreeding ranks does not imply a correct estimation of the “absolute” value of the inbreeding coefficient. Indeed, with both softwares and populations, F_{ROH} and F_{HBD} are constantly slightly above F_{IBD} (estimated with segments coalescing less than 100 time-steps ago) even when the SNP density increases (Figure 2a–d). PLINK did not detect ROHs with SNP densities below 10 SNPs/Mb, resulting in F_{ROH} of zero for all individuals (Figure S1a). RZooRoH always detected HBD segments, independently of the SNP density we tested. The variance among subsampling replicates was large and the rank of individuals' inbreeding was poorly conserved for SNP densities below 1 SNPs/Mb (Figure 2b,d; Figure S1b). We show in the Supporting Information that F_{HOM} , an estimator of inbreeding coefficient relying on the difference between the

observed and expected heterozygosity under Hardy–Weinberg yielded similar results to RZooRoH for the same SNPs densities (Figure S2; Table S3).

Concerning medium and high SNP densities (both SNP arrays and WGS), we see little effect of the sequencing method or the software used on F_{ROH} or F_{HBD} estimation: inbreeding coefficients estimates were always consistent with F_{IBD} (Figure 3, $r^2 > 0.97$). All sequencing methods resulted in slightly higher inbreeding coefficients, especially both arrays, but the rank of inbreeding was always conserved among individuals (Figure 3; Table S3). Interestingly, with PLINK and at similar densities a homogeneous spacing between SNPs (the SNP arrays) resulted in better correlations with F_{IBD} compared to RAD-sequencing in the small population (Figure 3a,c; Table S3).

3.2 | Distribution of ROHs and HBD segments

Figure 4 shows distribution of ROHs and HBD segments among the different length classes as the mean per individual (among simulation and subsampling replicates) total ROHs or HBD segments length falling within each length class. Horizontal black lines represent our

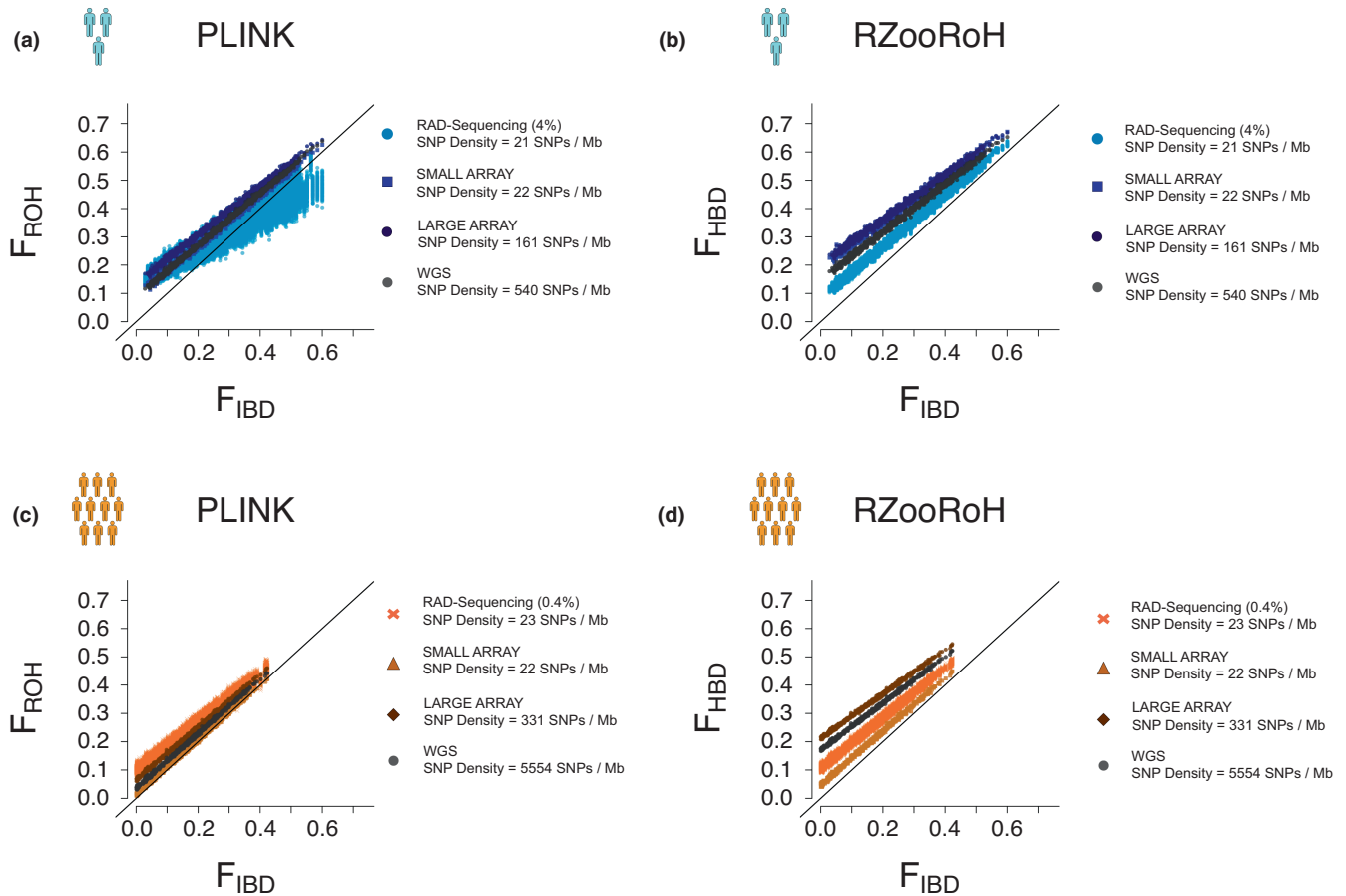


FIGURE 3 Comparison between runs of homozygosity (ROH) or homozygous-by-descent (HBD) estimated with different sequencing methods on the y-axis and identical-by-descent (IBD) segments (the true fraction of genome within IBD segments coalescing less than 100 time-steps ago) on the x-axis. Each point represents one individual (for one subsampling replicate within one simulation replicate). The black line represents the equality line ($x = y$). Blue points represent individuals from the small population and orange from the large population. Within these two colour categories, a change in shade represents an increase in single nucleotide polymorphism (SNP) density (fraction of genome subsampled indicated between the parentheses for RAD-sequencing). (a) Small population; ROHs were called with PLINK. Please note that points for the small array and whole-genome sequencing (WGS) perfectly overlap. (b) Small population; HBD segments were called with RZooRoH. Please note that points for both arrays perfectly overlap. (c) Large population; ROHs were called with PLINK. Please note that points for the large array and WGS (almost) perfectly overlap. (d) Large population; HBD segments were called with RZooRoH.

gold standard: the true mean (among simulation replicate) individual total IBD segment lengths estimated from simulated data for each length class. Bar plots represent the mean (among simulation and subsampling replicate) difference between the estimated distributions (ROHs or HBD segments) and the truth (IBD segments). Thus, bar plots above the horizontal black segment indicate an overestimation while bar plots below the segment an underestimation. In addition, the y-axis starts at 0 indicating that no IBD segments of the particular length class has been detected if the bar reaches the bottom of this axis. Compared to the true distribution of IBD segments (relative to a reference population from 100 reproductive cycles ago), almost all sequencing methods and both softwares resulted in higher mean length of ROHs or HBD segments falling into the smaller length class (Figure 4). At similar SNPs densities (i.e. with both SNP arrays), this overestimation was stronger for RZooRoH compared to PLINK, especially in the small population. However, there is no such overestimation in the small population when we

compare these distributions to older true IBD segments which coalesced less than 1000 reproductive cycles ago suggesting that these segments are not wrongly identified, they simply come from older coalescence events (Figure S3).

With WGS, we can correctly estimate distribution of ROHs and HBD segments larger than 2 Mb (Figure 4). In addition, both SNP arrays, allowed correct estimation of total lengths of ROHs and HBD segments larger than 4 Mb in the small populations and larger than 2 Mb in the large population (Figure 4). These results suggest that medium and high SNPs density datasets can be confidently used for ROHs and HBD segments detection with both PLINK and RZooRoH.

Concerning RAD-sequencing, PLINK allowed the correct estimation of distribution of ROHs with a SNP density around 22 SNPs/Mb (Figure 4a,c). On the other hand, RZooRoH yielded accurate distribution of HBD segments with a SNP density of 11 SNPs/Mb and seven SNPs/Mb in the small and large populations, respectively

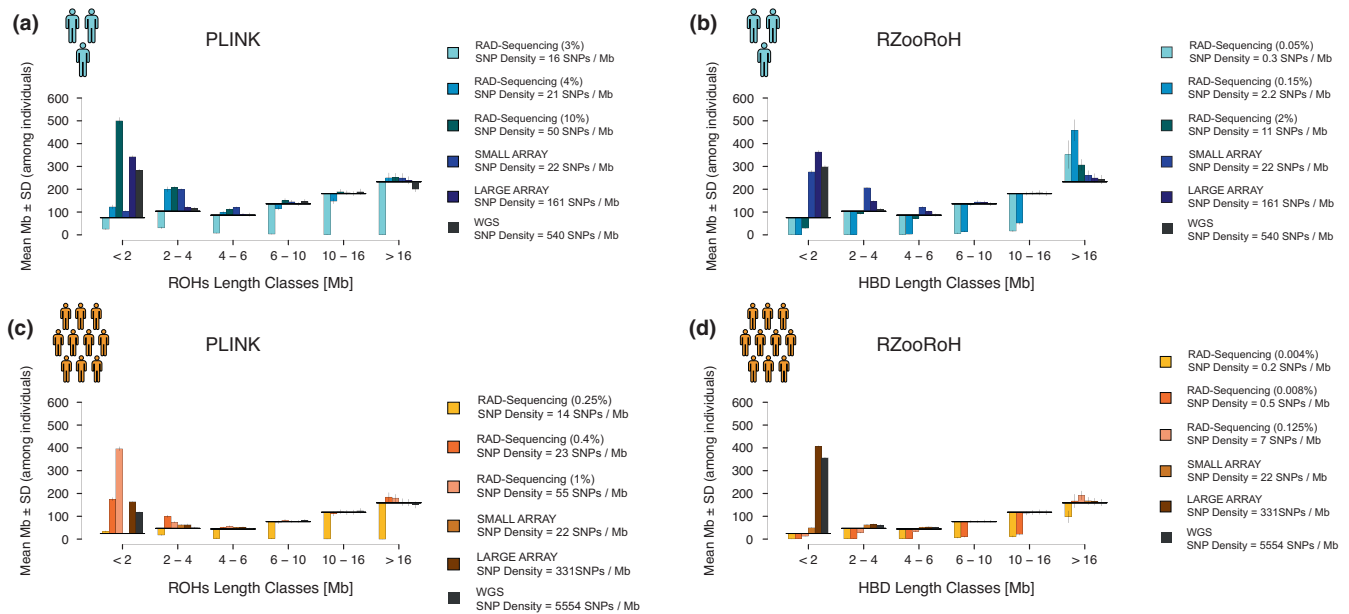


FIGURE 4 Comparison of distribution of runs of homozygosity (ROHs) (a and c) and homozygous-by-descent (HBD) segments (b and d) between the different sequencing methods and the true distributions of identical-by-descent (IBD) segments (defined as segments which coalesced less than 100 reproductive cycles ago). Black horizontal lines correspond to the total length of IBD segments per individual (y-axis) falling into the different length classes (x-axis). Bar plots show the mean (\pm SD) difference between the mean total length of IBD segment and their estimation (ROHs and HBD segments) for each sequencing method. Bar plots below the horizontal black line indicate an underestimation while bar plots above the horizontal black line indicate an overestimation of the total length of segments. Mean (\pm SD) are among individuals, simulation and subsampling replicates. (a) Distribution of ROHs from the small population; ROHs were called with PLINK. (b) Distribution of HBD segments from the small population; HBD segments were called with RZooRoH. (c) Distribution of ROHs from the large population; ROHs were called with PLINK. (d) Distribution of HBD segments from the large population; HBD segments were called with RZooRoH

(Figure 4b,d). For lower SNP densities, RZooRoH tended to merge small adjacent HBD segments into larger ones (Figure 4d).

We also investigated ROHs calling accuracy with *PLINK* and the default parameters for all reduced genomic representations. We show in the Supporting Information that similarly to what was observed in Figure 2, F_{ROH} can be correctly estimated with SNP densities higher than 22 SNPs/Mb (Figure S4 and Figure S5). However, the distribution of ROHs are always biased with a large overestimation of small IBD segments and underestimation of large IBD segments even with both SNP arrays (Figure S6). These results emphasize the importance of fine-tuning *PLINK* parameters when working with reduced genomic representations.

3.3 | Fraction of genome assigned within and outside IBD segments

Figure 5 shows the mean fraction of genome which has been correctly (true-non-IBD and true-IBD) and incorrectly (false-non-IBD and false-IBD) assigned to ROHs or HBD segments. Concerning RAD-sequencing, *PLINK* resulted in high fractions of genome incorrectly assigned to ROHs (false-IBD) (Figure 5a,c, left column). On the other hand, RZooRoH resulted in 90 and 95% of the genome correctly assigned within our outside IBD segments with 2% (SNP density = 11 SNPs/Mb) and 0.125% (SNP density = 7 SNPs/Mb) of

the genome sequenced in the small and large populations respectively (Figure 5b,d, right column). Concerning SNP arrays and WGS, the “incorrectly” assigned ROHs and HBD segments were mostly false positive (false-IBD). However, we show in Figure S7 that these false positive become true positive when compared to IBD segments which coalesced less than 1000 reproductive cycles ago. This indicates that these fragments come from coalescence events older than 100 reproductive cycles ago. In the large population, RZooRoH still resulted in a few false-positive when compared to distribution of IBD segments from less than 1000 reproductive cycles ago, with both the large array and WGS (Figure S7d).

3.4 | Cattle simulations

Figure 6 shows the comparison of F_{ROH} and F_{HBD} estimates, distribution of ROHs and HBD segments and fractions of genome correctly assigned within and outside IBD segments (coalescing less than 1000 reproductive cycles ago) for WGS and both SNP arrays in the cattle population. For this population, we consider IBD segments coalescing less than 1000 reproductive cycles ago because the major part of inbreeding comes from ancient coalescence events and both softwares resulted in poor IBD segments estimation when compared to segments coalescing less than 100 reproductive cycles ago (Figure S10). In this population,

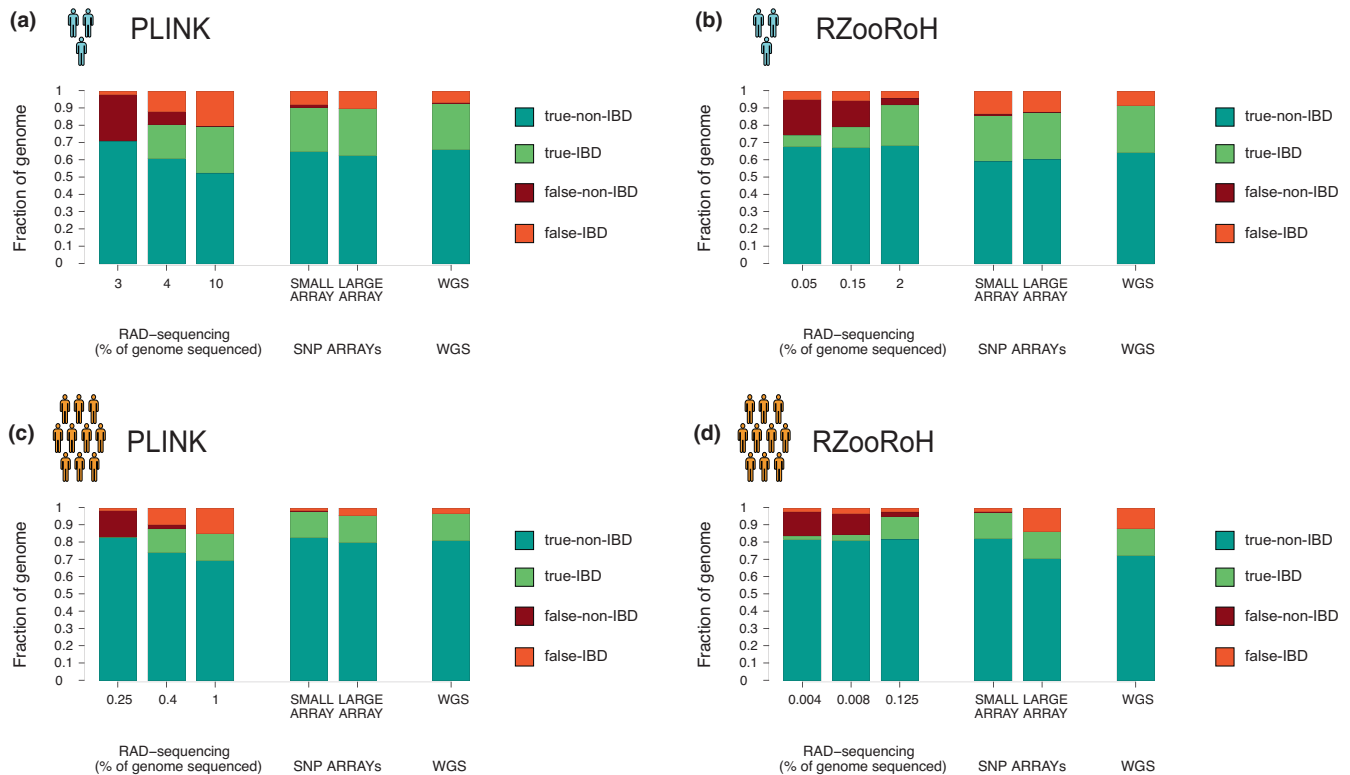


FIGURE 5 For each sequencing method, fraction of genome correctly assigned outside identical-by-descent (IBD) segments (true-non-IBD), correctly assigned within IBD segments (true-IBD), incorrectly assigned outside IBD segments (false-non-IBD) and incorrectly assigned within IBD segments (false-IBD) are represented. Values are averaged among individuals as well as both simulation and subsampling replicates. (a) Small population; runs of homozygosity (ROHs) were called with PLINK. (b) Small population; HBD segments were called with RZooRoH. (c) Large population; ROHs were called with PLINK. (d) Large population; homozygous-by-descent (HBD) segments were called with RZooRoH

PLINK resulted in perfect estimation of the inbreeding coefficient (Figure 6a; $r^2(F_{ROH\ WGS}, F_{IBD}) = 0.999$; $r^2(F_{ROH\ LARGE\ ARRAY}, F_{IBD}) = 0.993$; $r^2(F_{ROH\ SMALL\ ARRAY}, F_{IBD}) = 0.996$) but biased distribution of ROHs even with WGS (Figure 6c). Compared to the true distribution of IBD segments (coalescing less than 1000 reproductive cycles ago), PLINK identified higher numbers of ROHs smaller than 16 Mb (bar plots above the horizontal line) and fewer ROHs larger than 16 Mb (bar plots below the horizontal line) (Figure 6c). On the other hand with RZooRoH the correlation between F_{HBD} and F_{IBD} were slightly lower (Figure 6d: $r^2(F_{HBD\ WGS}, F_{IBD}) = 0.937$, $r^2(F_{HBD\ LARGE\ ARRAY}, F_{IBD}) = 0.983$; $r^2(F_{HBD\ SMALL\ ARRAY}, F_{IBD}) = 0.955$) but distribution of HBD segments were closer to the true distribution of IBD segments (Figure 6f). Similar to previous observations, RZooRoH detected fewer small HBD segments (<2Mb) with the small and to a lesser extend the large SNP array. Finally, concerning the fraction of genome correctly and incorrectly assigned to IBD segments, 83, 88 and 91% of the genome were correctly assigned within or outside ROHs with PLINK using the small array, large array and WGS, respectively (Figure 6b). With RZooRoH, 84, 87 and 85% of the genome were correctly assigned within or outside HBD segments with the small array, the large array and WGS (Figure 6e). With both arrays and softwares, the wrongly assigned fraction of genome was mainly false-negative (false-non-IBD) but both softwares resulted in a few false-positive (false-IBD) with WGS data (Figure 6b,e).

4 | DISCUSSION

4.1 | Summary

We investigated the capacity of WGS, SNP arrays and RAD-seq to perform IBD segments analyses using either the observational-based runs of homozygosity (ROHs) calling approach implemented in PLINK (Chang et al., 2015) or the model-based homozygous-by-descent (HBD) segments calling approach implemented in RZooRoH (Bertrand et al., 2019; Druet & Gautier, 2017). We show that both methods can be used with medium to high SNPs density data sets in simulations with constant population sizes and proportion of inbreeding. However, for RAD-seq, PLINK required a SNP density above 22 SNPs/Mb to keep the ranking among individuals F_{ROH} and correct estimates of distribution of ROHs. On the other hand, RZooRoH only required a SNP density of 11 SNPs/Mb to obtain correct distribution of HBD segments and the rank of individual inbreeding coefficients F_{HBD} was conserved when the SNP density was above 2 SNPs/Mb. We also show that in the cattle population, PLINK did not estimate distribution of IBD segments as accurately as RZooRoH even with WGS data. Finally, we show in Supporting Information that F_{HOM} , a non-IBD segments-based estimate of individual F , is as accurate as F_{HBD} estimated with the model-based approach for the same SNP density.

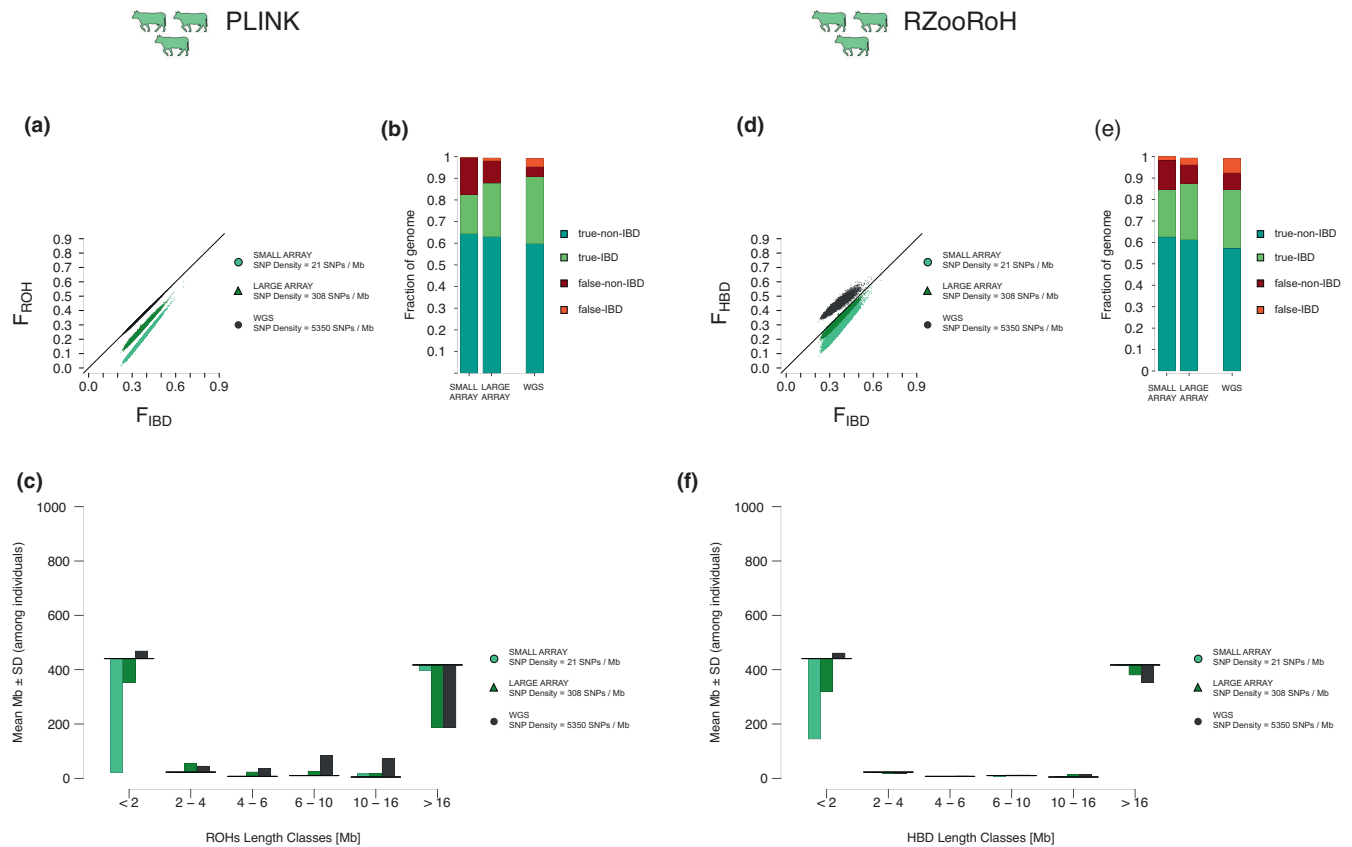


FIGURE 6 Comparison of runs of homozygosity (ROHs) and homozygous-by-descent (HBD) segment detection with whole-genome sequencing (WGS) and both single nucleotide polymorphism (SNP) arrays in the cattle population and for both identical-by-descent (IBD) segment detection methods. (a) Comparison of F_{ROH} estimated with PLINK on WGS data and both SNPs arrays with regard to the true fraction of genome within IBD segments (coalescing less than 1000 reproductive cycles ago): F_{IBD} . (b) Fraction of genome correctly assigned outside IBD segments (true-non-IBD), correctly assigned within IBD segments (true-IBD), incorrectly assigned outside IBD segments (false-non-IBD) and incorrectly assigned within IBD segments (false-IBD) for WGS and both SNP arrays. ROHs were called with PLINK. (c) Comparison of distribution of ROHs with WGS and both arrays with regard to the true distribution of IBD segments (coalescing less than 1000 generations ago). Horizontal black lines represent the true mean (among simulation replicates) individual total length of IBD segments estimated from simulated data for each length class. Bar plots represent the mean (among simulation and subsampling replicate) difference between the estimated distributions (ROHs) and the truth (IBD segments). ROHs were called with PLINK. (d) Comparison of F_{HBD} estimated with RZooRoH on WGS data and both SNP arrays with regard to the true fraction of genome within IBD segments (coalescing less than 1000 generations ago): F_{IBD} . (e) Fraction of genome correctly assigned outside IBD segments (true-non-IBD), correctly assigned within IBD segments (true-IBD), incorrectly assigned outside IBD segments (false-non-IBD) and incorrectly assigned within IBD segments (false-IBD) for WGS and both arrays. HBD segments were called with RZooRoH. (f) Comparison of distribution of HBD segments with WGS and both arrays with regard to the true distribution of IBD segments (coalescing less than 1000 generations ago). Horizontal black lines represent the true mean (among simulation replicate) individual total length of IBD segments estimated from simulated data for each length class. Bar plots represent the mean (among simulation and subsampling replicate) difference between the estimated distributions (HBD segments) and the truth (IBD segments). HBD segments were detected with RZooRoH

4.2 | PLINK vs. RZooRoH

With WGS and in the small and large populations, the two softwares yielded similar results concordant with the true IBD segments (extracted from simulated data as IBD segments which coalesced less than 100 reproductive cycles ago) for all segments larger than 2 Mb. We showed that both methods, but especially RZooRoH, identify a larger number of small IBD segments (<2 Mb) than the truth. However, we also show that these segments come from more ancient coalescence events. It makes sense that RZooRoH identifies a higher number of smaller HBD segments

compared to PLINK since it has no constraint on HBD minimum segments size. From the formula presented in Thompson (2013), the length of the IBD segment in centimorgans (cM) l is a simple function of the number of generations g since the coalescence event: $l = 100 / 2g$, our PLINK threshold corresponds to 0.1 cM (100 kb) in humans and thus coalescence events 500 reproduction events ago on average (with a very large variance [Speed & Balding, 2015]).

It is important to remember that even with WGS, IBD segment detection is still challenging. Beside the IBD segments detection method, many parameters can influence the result of IBD segments

calling. For instance, a minimum size threshold can be chosen. However, it has been shown with simulations that small segments can also result from recent coalescence events (Speed & Balding, 2015), thus neglecting smaller regions might lead to an underestimation of the inbreeding status of the individual or the population. Another important parameter is whether mutations (and sequencing errors) should be considered by allowing heterozygous SNPs in IBD segments and how many heterozygous markers are to be allowed can greatly differ among studies. To summarize, no consensus exist nowadays on which method and parameters are the best and further investigation is needed.

In this project, we used the true IBD values for parameter optimization with PLINK, not available with empirical data. We note that previous studies used WGS as the “gold-standard” and/or high coverage assembly data for parameters optimization (Duntsch et al., 2021; Meyermans et al., 2020; Mueller et al., 2022). Mueller et al., (2022) showed adapting parameters allow to get similar results between WGS and reduced genomic representations with smaller fractions of genomes: the authors used PLINK to call ROHs with RAD-sequencing, then tested ROHs calling with several different settings for three individuals for which they also had WGS data and extracted the settings which best conserved the rank of inbreeding found with WGS. We want to stress that WGS data might not be available or might not result in accurate IBD segments detection as observed with our cattle population. Finally, varying settings need to be done with caution: different settings can increase the number of ROHs detected but also the likelihood of noncorrect calls and maybe bias the individuals inbreeding ranking (Meyermans et al., 2020).

4.3 | Estimation of F_{ROH} and F_{HBD}

With WGS, both arrays and when a large portion of the genome is sequenced with RAD-sequencing, ranks of inbreeding were always conserved among individuals. Our results are consistent with Kardos et al. (2018) who showed (see Supporting Information) that F_{HBD} estimated with 10,000 loci, and a home-made script based on a likelihood ratio method (adapted from Pemberton et al., 2012), are similar to F_{HBD} estimated with WGS in an inbred wild wolf population.

With RAD-sequencing, accurate results of the fraction of genome within IBD segments depends on the fraction of genome sequenced, which drastically differ between both softwares. RZooRoH required a fraction of the genome 11 times smaller compared to PLINK for conserved ranking of inbreeding estimates among individuals. It is expected that model-based approaches perform better with low SNPs-densities as the distances between SNPs are taken into account in the model (Bertrand et al., 2019). Hence, we strongly recommend RZooRoH when working with reduced genomic data. Nevertheless, when the SNPs density is sufficient, the conservation of inbreeding ranks was observed with both softwares and is consistent with other studies: Duntsch et al. (2021) compared ROHs and HBD segments estimates from PLINK and

RZooRoH with RAD-sequencing, a custom-made array and WGS in few hihi (*Notiomystis cincta*; a nonmodel bird species) individuals. They found conserved individual inbreeding ranks with all reduced representations.

In the cattle population and with both SNP arrays, F_{ROH} and F_{HBD} estimates were lower than F_{IBD} . The missing portion of inbreeding was coming from small IBD segments with more ancient coalescence events. Indeed, recent inbreeding is easier to capture: with lower marker density you are not able to capture small segments (Druet & Gautier, 2017). This is supported by Sole et al. (2017) who compared SNP arrays of different sizes in cattle and showed that all arrays capture the same levels of recent inbreeding but higher densities allow older inbreeding to be captured.

Inbreeding coefficients are used for inbreeding depression studies, which are key for understanding the evolution of populations and for conservation managements in endangered species (Lynch & Walsh, 1998). If the rank of individuals' F is conserved, it ensures that the direction of the correlation between inbreeding estimates and phenotypes (the sign of the regression slope), and thus the general effect of inbreeding on the trait, is correctly estimated. However, underestimating the inbreeding coefficients can lead to an overestimation of the magnitude of inbreeding depression (the regression slope absolute value will be steeper). We showed in Supporting Information that F_{HOM} performed as well as model-based RZooRoH approach with similar SNPs densities. This last result is consistent with another study which showed that 5000 markers are sufficient to obtain genomic kinship values similar to pedigree estimates (Goudet et al., 2018). Hence, we strongly advise to use SNP independent measures (i.e., not based on IBD segments) or model-based HBD segments approaches for inbreeding studies when SNP density is low. However, it is important to keep in mind that F_{HOM} can only be used if the number of individuals is large enough to allow correct estimation of the allelic frequencies.

4.4 | Distribution of ROHs and HBD segments

Distribution of IBD segments are used to describe the demographic history of the populations (Ceballos et al., 2018; Kirin et al., 2010) as it is possible to link the length of an IBD segment to the number of generations back to the common ancestor (Thompson, 2013). Our study suggests that with both WGS and SNP arrays, length of ROHs and HBD segments values are correctly estimated. On the contrary, with RAD-sequencing, length of ROHs and HBD segments should not be trusted at low SNP densities, meaning that age estimation for these segments is currently impossible. In the cattle population and with WGS data, RZooRoH resulted in more accurate distribution of HBD segments compared to PLINK. This was not the case in the constant sizes population simulations. This difference is due to the more complex demographic history of the cattle population which underwent both a strong bottleneck and intense selective pressures. We hypothesize that the difference observed between both softwares is due to the way they handle

heterozygous SNPs. RZooRoH uses the probability of observing a heterozygous SNP in an HBD segment per base-pair, automatically adapting the number of heterozygous SNPs allowed in an HBD segment to its size. This is not the case with PLINK. Hence, manually increasing the maximum number of heterozygous SNPs allowed in a ROH will merge small adjacent ROHs belonging to the same IBD segment into a larger one; but might also increase the number of false positives. Consequently, we advise researchers to use model-based approaches when working with populations with complex demographic history.

As we expect all the bias mentioned above to be similar among populations, relative comparisons between populations genotyped with the same sequencing method lead to reliable results as shown by Kirin et al. (2010): these authors used the HGDP data set genotypes (from 2009) with Illumina 650Y product and PLINK and compared distribution of ROHs among various human populations. The populations undergoing contemporaneous inbreeding such as West and South Asians and Oceanians harboured a higher fraction of genome within long ROHs while populations from central America harboured lower numbers of large ROHs but higher numbers of small ROHs, consistent with an history of ancient small effective population size. Similarly, Mastrangelo et al. (2016) compared distribution of ROHs estimated with PLINK from three dairy cattle breeds genotyped with the “small array” (Illumina BovineSNP 50 beadchip) to compare the different breeds and to assess their inbreeding status. The authors found that Italian Holstein individuals harboured a high number of short ROHs suggesting that inbreeding in this breed is mostly caused by ancient relatedness within the population rather than recent mating between relatives. Individuals from two local breeds *Modicana* and *Cinisara* harboured a high number of large ROHs, suggesting recent mating events between relatives. The authors concluded that the implementation of a monitored breeding programme aimed at reduced consanguinity was necessary in these local breeds.

With WGS, both SNP arrays and high SNPs density RAD-sequencing, we obtained correct assignment of the genome within and outside IBD segments above 90%. The remaining incorrectly assigned regions correspond mostly to IBD segments with older coalescence times or segments in between adjacent IBD segments. This trend indicates that IBD regions are correctly detected with medium to high densities and thus suggests that they can be confidently used for homozygosity mapping studies.

4.5 | Effect of population size

Population size influenced the minimum proportion of genome required: the larger population required a smaller fraction of genome to obtain accurate IBD estimation. However, the larger population harboured higher genetic variation and thus a higher number of SNPs. Consequently, we were more likely to have SNPs in the subsampled regions with RAD-sequencing. Since the simulations scheme forced mating between related individuals and since we performed a random stratified individual sampling independently

of the effective size of the population, we do not expect distribution of IBD segments to reflect both populations “true” distributions. Consequently, in this study we cannot quantify the true effect of population size (other than the number of SNPs) on IBD segments detection. To disentangle these effects, one could vary the mutation rate so that both populations have similar SNP densities.

4.6 | Limitations

We want to stress that the fractions of the genome sequenced used in the present study are indicative and do not correspond to “true” proportion of genome sequenced needed to obtain meaningful results. They come from simulated data and this fraction will be lower after quality filtering with empirical data, especially for RAD-sequencing and WGS where genotyping rates are lower. In addition, we did not include the effect of allele dropout when we simulated RAD-sequencing, which could influence the accuracy of IBD segments detection. Indeed, individuals heterozygous at restriction sites might appear homozygous which can bias further population genetic inferences, especially in large populations (Arnold et al., 2013; Gautier et al., 2013). Finally, we did not include in this project the method for IBD segments detection based on the odds ratio comparison of the likelihood of the genotype if autozygous or allozygous (Broman & Weber, 1999; Pemberton et al., 2012) but we expect that this method can handle reduced genomic representations since it takes linkage disequilibrium into account.

4.7 | Conclusion

Using simulated data, we compared F_{ROH} and F_{HBD} estimates as well as distribution of ROHs and HBD segments to the true IBD (from simulated data and calculated as segments which coalesced less than 100 reproductive cycles ago). F_{ROH} and F_{HBD} can be correctly estimated with all sequencing methods when the SNP density is above 22 SNPs per Mb with PLINK and above two SNPs per Mb with RZooRoH. With lower SNPs densities, F_{HOM} , a genomic estimate of inbreeding coefficients not based on ROHs, is as accurate as the model-based estimate, for a fraction of the computing time. We would therefore recommend using independent SNPs-based genomic estimates such as F_{HOM} for inbreeding quantification with reduced genomic representation, unless the number of individuals analysed is too small to allow a correct estimation of the population alleles frequency or mean coancestry. Regarding distribution of ROHs and HBD segments, even though the majority of the genome is correctly assigned within and outside IBD segments, both softwares failed to capture IBD segments with older coalescence times at low densities. This still allows comparing populations analysed with the same methodology but prevents comparing distribution of ROHs or HBD segments from studies using different reduced genomic methods. In addition, in a population with a more complex demographic history, only RZooRoH resulted in accurate

distribution of HBD segments with WGS data. To conclude, we find little advantages in using IBD segments-based estimates of inbreeding at low SNPs densities and show that only model-based approaches can be used for distribution of HBD segments quantification at such low SNP densities and in populations with complex demographic histories.

AUTHOR CONTRIBUTIONS

Eléonore Lavanchy and Jérôme Goudet conceptualized the study. Eléonore Lavanchy conducted the study, performed all simulations and analyses and drafted the manuscript. Eléonore Lavanchy and Jérôme Goudet wrote and revised the manuscript.

ACKNOWLEDGEMENTS

We are grateful to the Association Wallonne de l'Elevage (awé) for accepting to share their cattle pedigree with us and to Tom Druet who kindly asked them on our behalf and provided feedbacks on this manuscript. We also want to thank Marty Kardos for his comments on this manuscript. Finally, we want to thank the reviewers: Anna Santure and an anonymous reviewer for their great comments and feedbacks. This study was funded by the Swiss National Science Foundation with grant 31003A_179358 to JG. Open Access Funding provided by University of Lausanne.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

All scripts necessary to generate genomic data and perform analyses used in this study have been made available on GitHub: <https://github.com/EléonoreLavanchy/ROHsReducedRep>.

ORCID

Eléonore Lavanchy  <https://orcid.org/0000-0003-4951-9332>

Jérôme Goudet  <https://orcid.org/0000-0002-5318-7601>

REFERENCES

- Åkesson, M., Liberg, O., Sand, H., Wabakken, P., Bensch, S., & Flagstad, Ø. (2016). Genetic rescue in a severely inbred wolf population. *Molecular Ecology*, 25(19), 4745–4756. <https://doi.org/10.1111/mec.13797>
- Alemu, S. W., Kadri, N. K., Harland, C., Faux, P., Charlier, C., Caballero, A., & Druet, T. (2020). An evaluation of inbreeding measures using a whole-genome sequenced cattle pedigree. *Heredity*, 126, 410–423. <https://doi.org/10.1038/s41437-020-00383-9>
- Alkuraya, F. S. (2013). The application of next-generation sequencing in the autozygosity mapping of human recessive diseases. *Human Genetics*, 132(11), 1197–1211. <https://doi.org/10.1007/s00439-013-1344-x>
- Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, 17(2), 81–92. <https://doi.org/10.1038/nrg.2015.28>
- Arnold, B., Corbett-Detig, R. B., Hartl, D., & Bomblies, K. (2013). RADseq underestimates diversity and introduces genealogical biases due to nonrandom haplotype sampling. *Molecular Ecology*, 22(11), 3179–3190. <https://doi.org/10.1111/mec.12276>
- Bertrand, A. R., Kadri, N. K., Flori, L., Gautier, M., & Druet, T. (2019). RZooRoH: An R package to characterize individual genomic autozygosity and identify homozygous-by-descent segments. *Methods in Ecology and Evolution*, 10(6), 860–866. <https://doi.org/10.1111/2041-210X.13167>
- Bittles, A. H., & Black, M. L. (2010). Consanguinity, human evolution, and complex diseases. *Proceedings of the National Academy of Sciences*, 107(suppl 1), 1779–1786. <https://doi.org/10.1073/pnas.0906079106>
- Bjelland, D. W., Weigel, K. A., Vukasinovic, N., & Nkrumah, J. D. (2013). Evaluation of inbreeding depression in Holstein cattle using whole-genome SNP markers and alternative measures of genomic inbreeding. *Journal of Dairy Science*, 96(7), 4697–4706. <https://doi.org/10.3168/jds.2012-6435>
- Bosse, M., Megens, H.-J., Madsen, O., Paudel, Y., Frantz, L. A. F., Schook, L. B., Crooijmans, R. P. M. A., & Groenen, M. A. M. (2012). Regions of homozygosity in the porcine genome: Consequence of demography and the recombination landscape. *PLoS Genetics*, 8(11), e1003100. <https://doi.org/10.1371/journal.pgen.1003100>
- Broman, K. W., & Weber, J. L. (1999). Long homozygous chromosomal segments in reference families from the Centre d'Étude du Polymorphisme Humain. *The American Journal of Human Genetics*, 65(6), 1493–1500. <https://doi.org/10.1086/302661>
- Browning, S. R., & Browning, B. L. (2010). High-resolution detection of identity by descent in unrelated individuals. *The American Journal of Human Genetics*, 86(4), 526–539. <https://doi.org/10.1016/j.ajhg.2010.02.021>
- Caballero, A., Villanueva, B., & Druet, T. (2020). On the estimation of inbreeding depression using different measures of inbreeding from molecular markers. *Evolutionary Applications*, 14, 416–428. <https://doi.org/10.1111/eva.13126>
- Carothers, A. D., Rudan, I., Kolcic, I., Polasek, O., Hayward, C., Wright, A. F., Campbell, H., Teague, P., Hastie, N. D., & Weber, J. L. (2006). Estimating human inbreeding coefficients: Comparison of genealogical and marker heterozygosity approaches. *Annals of Human Genetics*, 70(5), 666–676. <https://doi.org/10.1111/j.1469-1809.2006.00263.x>
- Ceballos, F. C., Joshi, P. K., Clark, D. W., Ramsay, M., & Wilson, J. F. (2018). Runs of homozygosity: Windows into population history and trait architecture. *Nature Reviews Genetics*, 19(4), 220–234. <https://doi.org/10.1038/nrg.2017.109>
- Chadeau-Hyam, M., Hoggart, C. J., O'Reilly, P. F., Whittaker, J. C., De Iorio, M., & Balding, D. J. (2008). Fregene: Simulation of realistic sequence-level data in populations and ascertained samples. *BMC Bioinformatics*, 9(1), 364. <https://doi.org/10.1186/1471-2105-9-364>
- Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, 4(1), 7. <https://doi.org/10.1186/s13742-015-0047-8>
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>
- de Jong, J. F., van Hooft, P., Megens, H.-J., Crooijmans, R. P. M. A., de Groot, G. A., Pemberton, J. M., Huisman, J., Bartoš, L., Iacolina, L., van Wieren, S. E., Ydenberg, R. C., & Prins, H. H. T. (2020). Fragmentation and translocation distort the genetic landscape of ungulates: Red Deer in The Netherlands. *Frontiers in Ecology and Evolution*, 8, 365. <https://doi.org/10.3389/fevo.2020.535715>
- Druet, T., & Gautier, M. (2017). A model-based approach to characterize individual inbreeding at both global and local genomic scales. *Molecular Ecology*, 26(20), 5820–5841. <https://doi.org/10.1111/mec.14324>

- Druet, T., & Gautier, M. (2022). A hidden Markov model to estimate homozygous-by-descent probabilities associated with nested layers of ancestors. *Theoretical Population Biology*, 145, 38–51. <https://doi.org/10.1016/j.tpb.2022.03.001>
- Duntsch, L., Whibley, A., Brekke, P., Ewen, J. G., & Santure, A. W. (2021). Genomic data of different resolutions reveal consistent inbreeding estimates but contrasting homozygosity landscapes for the threatened Aotearoa New Zealand hihi. *Molecular Ecology*, 30(23), 6006–6020. <https://doi.org/10.1111/mec.16068>
- Forutan, M., Ansari Mahyari, S., Baes, C., Melzer, N., Schenkel, F. S., & Sargolzaei, M. (2018). Inbreeding and runs of homozygosity before and after genomic selection in north American Holstein cattle. *BMC Genomics*, 19(1), 98. <https://doi.org/10.1186/s12864-018-4453-z>
- Forutan, M., Mahyari, S. A., Baes, C., Melzer, N., Schenkel, F. S., & Sargolzaei, M. (2018). Inbreeding and runs of homozygosity before and after genomic selection in north American Holstein cattle. *BMC Genomics*, 19(1), 98. <https://doi.org/10.1186/s12864-018-4453-z>
- Franklin, I. R. (1977). The distribution of the proportion of the genome which is homozygous by descent in inbred individuals. *Theoretical Population Biology*, 11(1), 60–80. [https://doi.org/10.1016/0040-5809\(77\)90007-7](https://doi.org/10.1016/0040-5809(77)90007-7)
- Frantz, L. A. F., Bradley, D. G., Larson, G., & Orlando, L. (2020). Animal domestication in the era of ancient genomics. *Nature Reviews Genetics*, 21(8), 449–460. <https://doi.org/10.1038/s41576-020-0225-0>
- Gautier, M., Gharbi, K., Cezard, T., Foucaud, J., Kerdelhué, C., Pudlo, P., Cornuet, J.-M., & Estoup, A. (2013). The effect of RAD allele dropout on the estimation of genetic variation within and between populations. *Molecular Ecology*, 22(11), 3165–3178. <https://doi.org/10.1111/mec.12089>
- Ghani, M., Reitz, C., Cheng, R., Vardarajan, B. N., Jun, G., Sato, C., Naj, A., Rajbhandary, R., Wang, L.-S., Valladares, O., Lin, C.-F., Larson, E. B., Graff-Radford, N. R., Evans, D., De Jager, P. L., Crane, P. K., Buxbaum, J. D., Murrell, J. R., Raj, T., ... Genetics, A. D. (2015). Association of Long Runs of homozygosity with Alzheimer disease among African American individuals. *JAMA Neurology*, 72(11), 1313–1323. <https://doi.org/10.1001/jamaneurol.2015.1700>
- Gibson, J., Morton, N. E., & Collins, A. (2006). Extended tracts of homozygosity in outbred human populations. *Human Molecular Genetics*, 15(5), 789–795. <https://doi.org/10.1093/hmg/ddi493>
- Goudet, J., Kay, T., & Weir, B. S. (2018). How to estimate kinship. *Molecular Ecology*, 27(20), 4121–4135. <https://doi.org/10.1111/mec.14833>
- Haller, B. C., & Messer, P. W. (2019). SLiM 3: Forward genetic simulations beyond the Wright–fisher model. *Molecular Biology and Evolution*, 36(3), 632–637. <https://doi.org/10.1093/molbev/msy228>
- Hildebrandt, F., Heeringa, S. F., Rüschenhoff, F., Attanasio, M., Nürnberg, G., Becker, C., Seelow, D., Huebner, N., Chernin, G., Vlangos, C. N., Zhou, W., O'Toole, J. F., Hoskins, B. E., Wolf, M. T. F., Hinkes, B. G., Chaib, H., Ashraf, S., Schoeb, D. S., Ovunc, B., ... Otto, E. A. (2009). A systematic approach to mapping recessive disease genes in individuals from outbred populations. *PLoS Genetics*, 5(1), e1000353. <https://doi.org/10.1371/journal.pgen.1000353>
- Hill, W. G., & Weir, B. S. (2011). Variation in actual relationship as a consequence of Mendelian sampling and linkage. *Genetics Research*, 93(1), 47–64. <https://doi.org/10.1017/S0016672310000480>
- Howrigan, D. P., Simonson, M. A., & Keller, M. C. (2011). Detecting autozygosity through runs of homozygosity: A comparison of three autozygosity detection algorithms. *BMC Genomics*, 12(1), 460. <https://doi.org/10.1186/1471-2164-12-460>
- Huisman, J. (2017). Pedigree reconstruction from SNP data: Parentage assignment, sibship clustering and beyond. *Molecular Ecology Resources*, 17(5), 1009–1024. <https://doi.org/10.1111/1755-0998.12665>
- Huisman, J., Kruuk, L. E. B., Ellis, P. A., Clutton-Brock, T., & Pemberton, J. M. (2016). Inbreeding depression across the lifespan in a wild mammal population. *Proceedings of the National Academy of Sciences*, 113(13), 3585–3590. <https://doi.org/10.1073/pnas.1518046113>
- Jones, O. R., & Wang, J. (2010). COLONY: A program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources*, 10(3), 551–555. <https://doi.org/10.1111/j.1755-0998.2009.02787.x>
- Kardos, M., Akesson, M., Fountain, T., Flagstad, O., Liberg, O., Olason, P., Sand, H., Wabakken, P., Wikenros, C., & Ellegren, H. (2018). Genomic consequences of intensive inbreeding in an isolated wolf population. *Nature Ecology & Evolution*, 2(1), 124–131. <https://doi.org/10.1038/s41559-017-0375-4>
- Kardos, M., Luikart, G., & Allendorf, F. W. (2015). Measuring individual inbreeding in the age of genomics: Marker-based measures are better than pedigrees. *Heredity*, 115(1), 63–72. <https://doi.org/10.1038/hdy.2015.17>
- Kariyat, R. R., & Stephenson, A. G. (2019). Inbreeding depression: It's not just for population biologists. *American Journal of Botany*, 106(3), 331–333. <https://doi.org/10.1002/ajb2.1256>
- Kelleher, J., Etheridge, A. M., & McVean, G. (2016). Efficient coalescent simulation and genealogical analysis for large sample sizes. *PLoS Computational Biology*, 12(5), e1004842. <https://doi.org/10.1371/journal.pcbi.1004842>
- Keller, L. F., & Waller, D. M. (2002). Inbreeding effects in wild populations. *Trends in Ecology & Evolution*, 17(5), 230–241. [https://doi.org/10.1016/S0169-5347\(02\)02489-8](https://doi.org/10.1016/S0169-5347(02)02489-8)
- Keller, M. C., Simonson, M. A., Ripke, S., Neale, B. M., Gejman, P. V., Howrigan, D. P., Lee, S. H., Lencz, T., Levinson, D. F., Sullivan, P. F., & The Schizophrenia Psychiatric Genome-Wide Association Study Consortium. (2012). Runs of homozygosity implicate Autozygosity as a schizophrenia risk factor. *PLoS Genetics*, 8(4), e1002656. <https://doi.org/10.1371/journal.pgen.1002656>
- Keller, M. C., Visscher, P. M., & Goddard, M. E. (2011). Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. *Genetics*, 189(1), 237–249. <https://doi.org/10.1534/genetics.111.130922>
- Kim, E.-S., Cole, J. B., Huson, H., Wiggans, G. R., Van Tassell, C. P., Crooker, B. A., Liu, G., Da, Y., & Sonstegard, T. S. (2013). Effect of artificial selection on runs of homozygosity in U.S. Holstein cattle. *PLoS One*, 8(11), e80813. <https://doi.org/10.1371/journal.pone.0080813>
- Kirin, M., McQuillan, R., Franklin, C. S., Campbell, H., McKeigue, P. M., & Wilson, J. F. (2010). Genomic runs of homozygosity record population history and consanguinity. *PLoS One*, 5(11), e13996. <https://doi.org/10.1371/journal.pone.0013996>
- Lencz, T., Lambert, C., DeRosse, P., Burdick, K. E., Morgan, T. V., Kane, J. M., Kucherlapati, R., & Malhotra, A. K. (2007). Runs of homozygosity reveal highly penetrant recessive loci in schizophrenia. *Proceedings of the National Academy of Sciences*, 104(50), 19942–19947. <https://doi.org/10.1073/pnas.0710021104>
- Leutenegger, A.-L., Prum, B., Génin, E., Verny, C., Lemaître, A., Clerget-Darpoux, F., & Thompson, E. A. (2003). Estimation of the inbreeding coefficient through use of genomic data. *The American Journal of Human Genetics*, 73(3), 516–523. <https://doi.org/10.1086/378207>
- Liu, L., Bosse, M., Megens, H.-J., de Visser, M., Groenen, M. A. M., & Madsen, O. (2020). Genetic consequences of long-term small effective population size in the critically endangered pygmy hog. *Evolutionary Applications*, 14(3), 710–720. <https://doi.org/10.1111/eva.13150>
- Lynch, M., & Walsh, B. (1998). *Genetics and analysis of quantitative traits* (Vol. 1). Sinauer Associates.
- Mastrangelo, S., Tolone, M., Di Gerlando, R., Fontanesi, L., Sardina, M. T., & Portolano, B. (2016). Genomic inbreeding estimation in small populations: Evaluation of runs of homozygosity in three local dairy cattle breeds. *Animal*, 10(5), 746–754. <https://doi.org/10.1017/S1751731115002943>
- McQuillan, R., Leutenegger, A.-L., Abdel-Rahman, R., Franklin, C. S., Pericic, M., Barac-Lauc, L., Smolej-Narancic, N., Janicijevic, B., Polasek, O., Tenesa, A., MacLeod, A. K., Farrington, S. M., Rudan, P.,

- Hayward, C., Vitart, V., Rudan, I., Wild, S. H., Dunlop, M. G., Wright, A. F., ... Wilson, J. F. (2008). Runs of homozygosity in European populations. *The American Journal of Human Genetics*, 83(3), 359–372. <https://doi.org/10.1016/j.ajhg.2008.08.007>
- Menges, E. S. (1991). Seed germination percentage increases with population size in a fragmented prairie species. *Conservation Biology*, 5(2), 158–164. <https://doi.org/10.1111/j.1523-1739.1991.tb00120.x>
- Meyermans, R., Gorssen, W., Buys, N., & Janssens, S. (2020). How to study runs of homozygosity using PLINK? A guide for analyzing medium density SNP data in livestock and pet species. *BMC Genomics*, 21(1), 94. <https://doi.org/10.1186/s12864-020-6463-x>
- Mueller, S. A., Prost, S., Anders, O., Breitenmoser-Würsten, C., Kleven, O., Klinga, P., Konec, M., Kopatz, A., Krojerová-Prokešová, J., Middelhoff, T. L., Obexer-Ruff, G., Reiners, T. E., Schmidt, K., Sindičić, M., Skrbinšek, T., Tām, B., Saveljev, A. P., Naranbaatar, G., & Nowak, C. (2022). Genome-wide diversity loss in reintroduced Eurasian lynx populations urges immediate conservation management. *Biological Conservation*, 266, 109442. <https://doi.org/10.1016/j.biocon.2021.109442>
- Nachman, M. W., & Crowell, S. L. (2000). Estimate of the mutation rate per nucleotide in humans. *Genetics*, 156(1), 297–304.
- Nalls, M. A., Guerreiro, R. J., Simon-Sanchez, J., Bras, J. T., Traynor, B. J., Gibbs, J. R., Launer, L., Hardy, J., & Singleton, A. B. (2009). Extended tracts of homozygosity identify novel candidate genes associated with late-onset Alzheimer's disease. *Neurogenetics*, 10(3), 183–190. <https://doi.org/10.1007/s10048-009-0182-4>
- Narasimhan, V., Danecek, P., Scally, A., Xue, Y., Tyler-Smith, C., & Durbin, R. (2016). BCFtools/RoH: A hidden Markov model approach for detecting autozygosity from next-generation sequencing data. *Bioinformatics*, 32(11), 1749–1751. <https://doi.org/10.1093/bioinformatics/btw044>
- Nietlisbach, P., Muff, S., Reid, J. M., Whitlock, M. C., & Keller, L. F. (2019). Nonequivalent lethal equivalents: Models and inbreeding metrics for unbiased estimation of inbreeding load. *Evolutionary Applications*, 12(2), 266–279. <https://doi.org/10.1111/eva.12713>
- Pemberton, T. J., Absher, D., Feldman, M. W., Myers, R. M., Rosenberg, N. A., & Li, J. Z. (2012). Genomic patterns of homozygosity in world-wide human populations. *The American Journal of Human Genetics*, 91(2), 275–292. <https://doi.org/10.1016/j.ajhg.2012.06.014>
- Peripolli, E., Munari, D. P., Silva, M. V. G. B., Lima, A. L. F., Irgang, R., & Baldi, F. (2017). Runs of homozygosity: Current knowledge and applications in livestock. *Animal Genetics*, 48(3), 255–271. <https://doi.org/10.1111/age.12526>
- Peripolli, E., Stafuzza, N. B., Munari, D. P., Lima, A. L. F., Irgang, R., Machado, M. A., Panetto, J. C., do, C., Ventura, R. V., Baldi, F., & da Silva, M. V. G. B. (2018). Assessment of runs of homozygosity islands and estimates of genomic inbreeding in Gyr (Bos indicus) dairy cattle. *BMC Genomics*, 19(1), 34. <https://doi.org/10.1186/s12864-017-4365-3>
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, 81(3), 559–575. <https://doi.org/10.1086/519795>
- Qanbari, S., & Wittenburg, D. (2020). Male recombination map of the autosomal genome in German Holstein. *Genetics Selection Evolution*, 52(1), 73. <https://doi.org/10.1186/s12711-020-00593-z>
- Quinlan, A. R. (2014). BEDTools: The Swiss-Army tool for genome feature analysis. *Current Protocols in Bioinformatics*, 47(1), 11.12.1–11.12.34. <https://doi.org/10.1002/0471250953.bi1112s47>
- Saremi, N. F., Supple, M. A., Byrne, A., Cahill, J. A., Coutinho, L. L., Dalén, L., Figueiró, H. V., Johnson, W. E., Milne, H. J., O'Brien, S. J., O'Connell, B., Onorato, D. P., Riley, S. P. D., Sikich, J. A., Stahler, D. R., Vilella, P. M. S., Vollmers, C., Wayne, R. K., Eizirik, E., ... Shapiro, B. (2019). Puma genomes from north and South America provide insights into the genomic consequences of inbreeding. *Nature Communications*, 10(1), 4769. <https://doi.org/10.1038/s41467-019-12741-1>
- Sole, M., Gori, A.-S., Faux, P., Bertrand, A., Farnir, F., Gautier, M., & Druet, T. (2017). Age-based partitioning of individual genomic inbreeding levels in Belgian blue cattle. *Genetics Selection Evolution*, 49, 92. <https://doi.org/10.1186/s12711-017-0370-x>
- Speed, D., & Balding, D. J. (2015). Relatedness in the post-genomic era: Is it still useful? *Nature Reviews Genetics*, 16(1), 33–44. <https://doi.org/10.1038/nrg3821>
- Stoffel, M. A., Johnston, S. E., Pilkington, J. G., & Pemberton, J. M. (2021). Genetic architecture and lifetime dynamics of inbreeding depression in a Wild mammal. *Nature Communications*, 12(1), 2972. <https://doi.org/10.1038/s41467-021-23222-9>
- Thompson, E. A. (2013). Identity by descent: Variation in meiosis, across genomes, and in populations. *Genetics*, 194(2), 301–326. <https://doi.org/10.1534/genetics.112.148825>
- Wang, J. (2016). Pedigrees or markers: Which are better in estimating relatedness and inbreeding coefficient? *Theoretical Population Biology*, 107, 4–13. <https://doi.org/10.1016/j.tpb.2015.08.006>
- Wang, S., Haynes, C., Barany, F., & Ott, J. (2009). Genome-wide autozygosity mapping in human populations. *Genetic Epidemiology*, 33(2), 172–180. <https://doi.org/10.1002/gepi.20344>
- Weir, B. S., & Goudet, J. (2017). A unified characterization of population structure and relatedness. *Genetics*, 206(4), 2085–2103. <https://doi.org/10.1534/genetics.116.198424>
- Wright, S. (1922). Coefficients of inbreeding and relationship. *The American Naturalist*, 56(645), 330–338. <https://doi.org/10.1086/279872>
- Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: A tool for genome-wide complex trait analysis. *The American Journal of Human Genetics*, 88(1), 76–82. <https://doi.org/10.1016/j.ajhg.2010.11.011>
- Yengo, L., Zhu, Z., Wray, N. R., Weir, B. S., Yang, J., Robinson, M. R., & Visscher, P. M. (2017). Detection and quantification of inbreeding depression for complex traits from SNP data. *Proceedings of the National Academy of Sciences*, 114(32), 8602–8607. <https://doi.org/10.1073/pnas.1621096114>
- Zhang, C., Wang, P., Tang, D., Yang, Z., Lu, F., Qi, J., Tawari, N. R., Shang, Y., Li, C., & Huang, S. (2019). The genetic basis of inbreeding depression in potato. *Nature Genetics*, 51, 374–378. <https://doi.org/10.1038/s41588-018-0319-1>
- Zhang, Q., Calus, M. P. L., Guldbbrandtsen, B., Lund, M. S., & Sahana, G. (2015). Estimation of inbreeding using pedigree, 50k SNP chip genotypes and full sequence data in three cattle breeds. *BMC Genomics*, 16, 88. <https://doi.org/10.1186/s12863-015-0227-7>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Lavanchy, E., & Goudet, J. (2023).

Effect of reduced genomic representation on using runs of homozygosity for inbreeding characterization. *Molecular Ecology Resources*, 23, 787–802. <https://doi.org/10.1111/1755-0998.13755>

Chapter II: Detecting inbreeding depression in structured populations

Authors: Eléonore Lavanchy^{1,2} Bruce S. Weir³ and Jérôme Goudet^{1,2}

¹ Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland

² Swiss Institute of Bioinformatics, University of Lausanne, Lausanne, Switzerland

³ Department of Biostatistics, University of Washington, Seattle WA 98195, USA

Chapter status: re-submitted after the first round of revisions (asking for minor revisions) on the 12th of February 2024 to *Proceedings of the National Academy of Sciences (PNAS)*. Under revision.

Authors contribution: JG and BSW conceptualized the study. EL and JG conducted the study and performed all simulations and analyses. EL and JG drafted the manuscript. EL, JG and BSW wrote and revised the manuscript.

Detecting inbreeding depression in structured populations

Abstract

Measuring inbreeding and its consequences on fitness is central for many areas in biology including human genetics and the conservation of endangered species. However, there is no consensus on the best method, neither for quantification of inbreeding itself nor for the model to estimate its effect on specific traits. We simulated traits based on simulated genomes from a large pedigree and empirical whole-genome sequences of human data from populations with various sizes and structure (from the 1,000 Genomes project). We compare the ability of various inbreeding coefficients (F) to quantify the strength of inbreeding depression: allele-sharing, two versions of the correlation of uniting gametes which differ in the weight they attribute to each locus and two identical-by-descent segments-based estimators. We also compare two models: the standard linear model (LM) and a linear mixed model (LMM) including a genetic relatedness matrix (GRM) as random effect to account for the non-independence of observations. We find LMM give better results in scenarios with population or family structure. Within the LMM, we compare three different GRM and show that in homogeneous populations, there is little difference among the different F and GRM for inbreeding depression quantification. However, as soon as a strong population or family structure is present, the strength of inbreeding depression can be most efficiently estimated only if (i) the phenotypes are regressed on F based on a weighted version of the correlation of uniting gametes, giving more weight to common alleles and (ii) with the GRM obtained from an allele-sharing relatedness estimator.

Introduction

Inbreeding is the result of mating between relatives and is often associated with reduced fitness, a phenomenon called inbreeding depression (ID) and which was observed in many different species such as humans [43, 7], other animals [39, 18, 31, 17], and plants [51].

Many different methods have been developed for inbreeding quantification and there is no consensus on which one is the best [2, 6, 14, 36, 50, 52, 11]. The classical approach

was first proposed by Sewall Wright in 1922 and makes use of pedigrees (called hereafter F_{PED}) [47]. With the advances in sequencing technologies, genomic-based inbreeding coefficients (hereafter called F_{genomic}) have been developed. Among these, some coefficients rely on the comparison between observed and expected heterozygosity such as F_{HOM} [8, 40], the expected allele sharing between individuals such as F_{AS} [52] or on the correlation between uniting gametes such as F_{UNI} [48]. In addition to estimating the realized inbreeding coefficient and requiring no prior knowledge of the mating behavior of the population, these genomic estimates are simple and straightforward to compute and do not require whole-genome sequencing (WGS) data; a few thousand SNPs are usually sufficient for reliable inbreeding estimation in humans [14]. However, they also have a disadvantage: they usually require quadratic moments of allelic proportions (except for F_{AS}) which are complex functions of sample allelic frequencies and coancestry coefficients [52] and can lead to biased estimates. Another inbreeding coefficient was proposed by McQuillan *et al.* (2008): F_{ROH} uses runs of homozygosity (ROHs), long homozygous stretches as a proxy for IBD segments within individuals [33]. A model-based approach relying on hidden Markov models (HMM) has also been developed for detecting IBD segments [28] by identifying homozygous-by-descent (HBD) segments. This model is the basis for many other model-based IBD segments detection methods such as `BCFTOOLS` [35], `BEAGLE` [3] and `RZOOROH` [12]. The inbreeding coefficient estimated with these model-based approaches will be called F_{HBD} from now on. One advantage of these methods is that they are not sensitive to poorly estimated allele frequencies, particularly in terms of rank of inbreeding conservation (i.e. which individuals are the most and less inbred), which can be very valuable when only a few individuals are available. However, it has been shown that these coefficients, and especially F_{ROH} , are sensitive to SNP density and parameters used, and there is no consensus on what is the most suitable set of parameters at present [34, 27].

How to quantify ID, although central to conservation genetics for decades [23, 11] (more details and references in the detailed introduction section in supplementary material), is still debated. This debate includes two sub-questions: which statistical model should be employed ? And which F ? Regarding the model, the classical approach consisted of using linear regression of the phenotypes on the inbreeding coefficient. However, other models have been utilized, such as Generalized Linear models (GLMs) with various link functions. In 2019, Nietlisbach *et al.* [36] compared different models and found that the common GLM models with logit link did not allow for accurate inbreeding depression

strength estimation. They propose using GLM with logarithm link functions. Ultimately, the type of model is largely dependent on the distribution of the trait.

Regarding the choice of which F is more accurate for quantifying ID, many studies have demonstrated that $F_{genomic}$ yields better results than F_{PED} [24, 4, 20, 15]. However, some studies found F_{UNI} to be more accurate than F_{ROH} [50], while others found that F_{ROH} provided the best estimates of ID [24, 20, 36, 42]. In 2020, Caballero *et al.* [6] used simulations and included several populations with different histories: they found that the optimal F actually depends on how large the population is. F_{ROH} did a better job at quantifying ID in populations with small effective size while F_{UNI} was better at predicting ID estimates in populations with large effective sizes. This result was later confirmed by Alemu *et al.* [2] who used SNP-array empirical cattle data for several groups of allelic frequencies and concluded that F_{UNI} and F_{GRM} (F_I and F_{III} respectively in [48]) are better at quantifying homozygosity at rare alleles while F_{ROH} and F_{HOM} are better for alleles at intermediate frequencies and correlate better with whole-genome homozygosity. Indeed, recessive deleterious alleles, which are thought to be responsible for inbreeding depression, should segregate at low frequencies in large populations as a result of negative selection. On the contrary, in small populations, drift can increase the frequency of deleterious recessive alleles to intermediate frequencies, making F_{ROH} and F_{HOM} more suitable for detecting ID. Indeed, in the simulations conducted by Yengo *et al.* [50], rare alleles always caused negative effects on fitness (referred to as DEMA, for Directional Effect of Minor Alleles). The authors showed that F_{HOM} (and thus F_{AS} since they have similar properties) is sensitive to DEMA while F_{UNI} and F_{ROH} are not. They also showed via simulations that all estimates of ID are somewhat sensitive to population structure, F_{UNI} being the least affected. They recommend estimating ID using Linkage Disequilibrium (LD) score and Minor Allele Frequency (MAF) bins, and summing the ID estimates from these bins as an overall estimate of ID for the trait.

In this paper we simulated traits based on both simulated and empirical WGS human data from populations with varying sizes and structure. We show that some F are more sensitive to population structure and DEMA than others. We confirm only some of Yengo *et al.* [50] results. Importantly, we show that accounting for the non-independence of observations with a mixed model via an allele-sharing based genomic relationship matrix (GRM) (rather than the standard GRM from GCTA) and using a modified version of F_{UNI}

which gives more weight to common alleles resolves most of the issues raised by Yengo *et al.* [50].

Material and Methods

All codes used in this manuscript can be found on [GitHub](#).

Simulated pedigrees

We simulated a polygamous pedigree from a dioecious population with overlapping generations (hereafter called PEDIGREE) using custom R scripts. The population started from 500 founders (equal numbers of males and females) and followed a polygamous mating system: female fertilities per time interval were drawn from a Poisson distribution with parameter $\lambda = 1$, mortality rate per time interval was set to 0.5, and only 10% of the males were allowed to reproduce at each time step. Matings were recorded for 25 time steps, resulting in a pedigree of 11,924 individuals (over 25 time steps).

In order to simulate the genotypes of the individuals, we proceeded in two steps. We used the `mspms` wrapper to the `msprime` software [22] to simulate the two haplotypes containing $L = 650,000$ loci for each founder individual. The L loci were uniformly distributed along a constant recombination map $20M$ long. For each reproduction event, the number of cross-overs was first drawn from a Poisson distribution and then randomly positioned along the genome. The non-founder genotypes were then obtained by drawing two gametes: one from each parent. For each gamete, the allele at the first locus is selected at random between the two alleles of the parent. The alleles at the next locus along the chromosome are copied from the chromosome with the chosen allele at the first locus until a recombination event occurs, at which point the alleles are copied from the other chromosome until the next crossing-over or the end of the chromosome.

To investigate the effect of using more realistic smaller sample sizes, we subsampled 2,500 individuals from the PEDIGREE population. We performed two types of subsampling: i) a random subsampling where individuals were subsampled completely randomly, ii) a stratified subsampling where we sought to retain the widest range of inbreeding coefficients in the subsampled population. Consequently, for this stratified subsampling individuals with $F_{UNI}^w \geq 0.2$ were always included and individuals with $F_{UNI}^w < 0.2$ were randomly selected until the population reached the desired size. 100 replicates

were performed for each subsampling. We also simulated smaller pedigrees (50, 100, 250 and 500 individuals) with lower numbers of founders (8, 16, 40, 80 respectively).

1,000 genomes

In order to extend our conclusions to smaller sample sizes and populations with stronger structure (which are common in wild and/or endangered species), we used empirical data from phase 3 from the 1,000 Genomes project [44]. We considered i) a small sample from a homogeneous population with a small effective size represented by 504 individuals from the super-population with East-Asian ancestry (EAS), ii) a small sample from a population with some admixture and larger effective population sizes represented by 661 individuals from the super-population with African ancestry and admixed individuals (AFR) and finally iii) a larger sample from a population with larger effective size and with strong genetic structure (global $F_{ST} = 0.083$) comprising all the 2,504 individuals (hereafter called WORLD) and represented by five super-populations: individuals with East-Asian ancestry (EAS), African ancestry (AFR), European ancestry (EUR), admixed American ancestry (AMR) and finally South-Asian ancestry (SAS). A more detailed description of the samples can be found at the [1,000 Genomes Project website](#). To extend our findings to even smaller sample sizes, we subsampled the WORLD populations to 50, 100, 250, and 500 individuals. In each subsampling, we ensured that the entire range of F was covered and that similar numbers of individuals were subsampled from each continent.

Simulated traits

We simulated traits based on equation 1 following [50]: we consider a trait y whose phenotype is partly determined by the genotypes at L_c causal loci with $h^2 = 0.8$. We assume these loci to be bi-allelic, with one allele encoding for an increase in the trait value (the plus allele) and the other encoding for a decrease in trait value (the minus allele). Dominance was also considered since inbreeding depression (ID) occurs only if there is directional dominance: when heterozygotes at loci encoding for the trait are closer on average to the homozygote for the plus allele [30]. If gene effects are purely additive or if dominance is not directional, there is no ID. Finally, we assume no epistasis between loci and no genotype-environment interaction.

For individual j , y_j is the individual trait value (its phenotype), calculated as the sum of allelic and genotypic effects over causal loci, an environmental effect and μ , the average trait value among all individuals. At locus l , x_{jl} is the minor allele count (MAC) $\in \{0, 1, 2\}$ of individual j . a_l represents the additive effect size of the alternate allele at locus l . d_l is the dominance effect size, the deviation of the heterozygous genotype from the mean of the two homozygotes. Finally, ϵ_j is the environmental contribution to the phenotype of individual j , drawn from a normal distribution.

$$y_j = \mu + \sum_{l=1}^{Lc} x_{jl} a_l + \sum_{l=1}^{Lc} x_{jl} (2 - x_{jl}) d_l + \epsilon_j \quad (1)$$

The strength of inbreeding depression b was set to -3 in all simulations, as in Yengo *et al.* [50]. The value corresponds to an average reduction in trait value of 0.75 standard deviation for an offspring resulting from a mating between siblings.

We used equation 1 to simulate traits with varying architectures. To avoid causal markers with extremely low frequencies, we first excluded loci with $MAF \leq 0.01$ for both the EAS and AFR populations and loci with $MAF \leq 0.001$ for both the PEDIGREE and WORLD populations. We then simulated traits using 1,000 randomly chosen SNPs (after MAF filtering). We initially drew both the raw (i.e. unscaled) additive effect sizes of the alternate allele and the raw dominance effect sizes from a uniform $[0, 1]$ distribution (other distributions were explored with almost no effect on the results (results not shown)). As we expect alleles causing ID to be counter selected and thus removed or maintained at a low frequency (proportionally to their detrimental effect), the raw effect sizes were scaled inversely to MAF $a_j = raw_{a_j}/p_j$ to mimic negative selection. We also scaled the dominance effects inversely to the locus expected heterozygosity $d_j = raw_{d_j}/(2p_j(1 - p_j))$. In addition, we attributed the same sign to the effect sizes of all minor alleles in order to include what Yengo *et al.* [50] called Directional Effect of Minor Alleles (DEMA) [50]. However, in order to investigate the effect of the parameters mentioned above, we also simulated traits where the additive and dominance effect sizes were left unchanged $a_j = raw_{a_j}$ and $d_j = raw_{d_j}$ and without DEMA. A summary of all the simulated scenarios can be found in table S1. In addition, graphical representation of the additive effect sizes and dominance coefficients distribution under these different scenarios can be found in figure S1.

Individuals inbreeding coefficients

We estimated individual inbreeding coefficients using several methods whose properties were recently described in detail in Zhang *et al.* [52]. Regarding the figures and tables presented in the main text, we do not filter on MAF for any of the F estimates. We use one allele-sharing-based estimator of inbreeding, hereafter called F_{AS} and described in [46, 52]:

$$F_{AS_j} = \frac{\sum_{l=1}^L (A_{jl} - A_{Sl})}{\sum_{l=1}^L (1 - A_{Sl})} \quad (2)$$

where A_{jl} indicates the identity of the two alleles an individual j carries at locus l : one for homozygous and 0 for heterozygous and A_{Sl} is the average allele sharing proportion at locus l for pairs of individuals $j, k, j \neq k$.

Then, we compare two versions of F_{UNI} (initially described in [48]) and which measure the correlation between uniting gametes. The first version (hereafter called F_{UNI}^u is the original F_{UNI} [48] measured as the average of ratios over SNPs (which attributes equal weight ($1/L$) to all loci and results in loci with rare alleles having larger influence on the estimated F):

$$F_{UNI_j}^u = \frac{1}{L} \sum_{l=1}^L \frac{x_{jl}^2 - (1 + 2p_l)x_{jl} + 2p_l^2}{2p_l(1 - p_l)} \quad (3)$$

Similarly to equation 1, $x_{jl} \in \{0, 1, 2\}$ is the MAC of individual j at locus l and p_l is the derived allele frequency at locus l .

The second version (hereafter called F_{UNI}^w is a modified version of F_{UNI} which measures the ratio of averages (rather than the average of ratios) and thus gives more weight to loci with larger expected heterozygosity (i.e. with MAF close to 0.5). We are not aware of other investigations using the ratio of averages estimator F_{UNI}^w in the context of ID estimation.

$$F_{UNI_j}^w = \frac{\sum_{l=1}^L (x_{jl}^2 - (1 + 2p_l)x_{jl} + 2p_l^2)}{\sum_{l=1}^L 2p_l(1 - p_l)} \quad (4)$$

We also used four Identical-by-descent (IBD) segments-based F . We identified runs of homozygosity (ROHs) with PLINK [40] and default parameters. We also modelled

Homozygous-by-descent (HBD) segments with `BCFTTOOLS` [35]. For both methods, we selected ROHs or HBD segments based on their size: either larger than 100Kb: $F_{\text{ROH.100KB}}$ and $F_{\text{HBD.100KB}}$ or larger than 1Mb: $F_{\text{ROH.1MB}}$ and $F_{\text{HBD.1MB}}$. For both methods the inbreeding coefficients were simply estimated as the fraction of genome falling within ROHs or HBD segments.

Finally, in the PEDIGREE population, we used the pedigree-based inbreeding coefficient: F_{PED} [47].

All inbreeding coefficients were estimated separately for each population of the 1,000 Genomes Project (EAS, AFR, WORLD) and using only the polymorphic SNPs in each population and population-specific allelic frequencies (for both F_{UNI}). Consequently, the same individual might have different F_{genomic} in the EAS and the WORLD population. This influenced only trivially the IBD segments-based inbreeding coefficients (F_{ROH} and F_{HBD}) but influenced greatly F_{AS} (though the rank of inbreeding among individuals was perfectly conserved) and both F_{UNI} (for which the rank of inbreeding among individuals was not conserved). Comparison among the different inbreeding coefficients per population can be found in supplementary material (figures S2 – S5). More details can be found in [52].

Estimation of inbreeding depression: b

We estimated the strength of ID (hereafter defined as b) using two different models. In the first model, b was estimated as the slope of regression of phenotypes on the different inbreeding coefficients with a classical linear model (LM):

$$\widehat{b}_{\text{LM}} = \text{Cov}(Y, F) / \text{Var}(F) \quad (5)$$

where Y is the vector of trait values and F is the vector of individual inbreeding coefficients estimates.

In the second model, we estimate b as the fixed effect coefficient associated with the inbreeding coefficient in the following linear mixed model (LMM):

$$Y = bX + \omega + \epsilon \quad (6)$$

where Y is the vector of trait values, X is a matrix with two columns, the first containing ones and the second the individual inbreeding coefficients, ω is the random component of the mixed model with $\omega \sim N(0, \tau K)$, K being the genomic relationship matrix (GRM) and τ the additive variance component. Finally, ϵ is the individual residual variance and is defined as $\epsilon \sim \sigma^2 I_n$. From this, b is estimated as follows:

$$\widehat{b}_{LMM} = (X'V^{-1}X)^{-1}X'V^{-1}Y \quad (7)$$

with $V = \tau K + \sigma^2 I_n$ [10]. We compared three GRMs we estimated using all loci (no MAF filtering). The first mixed model included a GRM derived from allele sharing [14], hereafter called LMMAS. We used the R `Hierfstat` [13] package to estimate K and the R `gaston` package [1] to estimate V and b . We could not use GCTA software to run the mixed model for this GRM because its leading eigenvalue is negative which the Choleski decomposition algorithm used for matrix inversion in GCTA cannot handle (it requires a positive definite matrix), while the Schur decomposition algorithm used in `gaston` can. We note that the standard GRM is not positive definite (one eigenvalue is 0), but the matrix to invert in the mixed model is not the GRM itself but $V = \tau K + \sigma^2 I_n$ which becomes positive definite and can be inverted if the heritability is smaller than one.

The second mixed model used the GCTA weighted GRM matrix [14, 45]. Similarly to F_{UNI}^w , this matrix uses the ratio of averages. For this model, we used GCTA and the R `SNPrelate` package to estimate the GRM. We then used the R `gaston` package for estimating V and b with the LMM.

Finally, the third mixed model used the GCTA unweighted GRM matrix [48] which (similarly to F_{UNI}^u) utilizes the average of ratios and thus gives equal weight to all loci. For this model, we used GCTA to estimate the GRM. We then estimated V and b with the LMM implemented in the R `gaston` package.

Note that the Average Information-Restricted Maximum Likelihood (AIREML) fitting method we used in the LMM is an iterative procedure and should result in unbiased estimates. In some cases, the model did not converge and gave highly biased b . For each scenario, regression model and population, the number of replicates which did not converge can be found in tables S7 – S9.

Application to an empirical dataset

A metapopulation of house sparrows (*Passer domesticus*) from several islands in Northern Norway has been monitored since 1993 and Niskanen *et al.* [38] investigated inbreeding depression on several traits and made available phenotype and genotype data on more than 3,100 adult individuals. The data set is ideal to illustrate our method as individuals belong to different islands, the data contains slight genetic structure and some individuals are highly related (see supplementary materials for further details).

We used only morphological phenotypes, as they can be analyzed with Linear Models. We removed information from non-autosomes (scaffold 32) but otherwise kept all SNPs for both F_{UNI}^W and the GRM_{AS} estimations to avoid biases when filtering for minor allele frequencies and LD [46]. We filtered out individuals who were not present as adults in one of the eight studied islands, as done in the original analysis [38]. The data set used for analysis contained 1,786 individuals genotyped at 181,529 SNPs. We compared the results of a simple LM with Sex and F_{UNI}^W as explanatory variables, to the LMM_{AS} model with Sex and F_{UNI}^W as fixed effects. We also present two additional linear mixed models: one with island and year nested in island as random effects, as done in the original article, and a “full” mixed model with all the random effects mentioned above. Linear models were carried out with the *lm* function of R, while the mixed models were carried out with the *lmer* function of the *lme4* package or the *lmm.aireml* function of the *gaston* package if the model contained a GRM. To test if b , the slope associated with F_{UNI}^W was significantly different from 0, we used the *score.fixed.linear* function of the *gaston* package.

Results

All the figures presented in the main text picture the scenario where alleles additive effect sizes and dominance coefficients are proportional to MAF and where there is a directional additive effect of minor alleles (DEMA) (i.e. the ADD & DOM & DEMA scenario from table S1) (see figure S1). The results for the other scenarios are shown and discussed in supplementary material (figures S10 – S17, tables S3 – S6).

Simulated pedigrees

Figure 1 presents the inbreeding depression (ID) strength estimates (b , see the methods section) for the different inbreeding coefficients (F), with two regression models in the PEDIGREE populations. The first column shows b estimated with the simple LM and the second column shows b estimated with LMM including the allele-sharing GRM as a random factor (LMM_{AS}). The first row shows results for the complete PEDIGREE population ($n = 11,924$). The second row shows results for a reduced sample size of the PEDIGREE population ($n = 2,500$, meant to match the size of the 1KG WORLD population) where subsampled individuals were chosen completely randomly. The third row also shows results for a reduced sample size of the PEDIGREE population ($n = 2,500$) but these individuals were selected to represent the entire spectrum of inbreeding values. The violin plots show b estimates distributions among the simulation replicates (100 replicates for the complete population, 10,000 replicates for both subsampled populations). The solid dark grey line is the true strength of ID ($b = -3$). The dashed red line represents the absence of ID ($b = 0$), indicating that ID was not detected in any replicate above this line. Root mean square error (RMSE) values associated with both models and populations are shown in table 1. Strikingly, in the PEDIGREE population, all F resulted in a biased estimation of b with the simple LM, whatever the sample size (figure 1, panels A, C and E; table 1). The inclusion of a GRM matrix as a random factor allowed for the correction of non-independence of observations and greatly improved b estimation (figure 1, panels B, D, and F; table 1). In the complete PEDIGREE population, we see little difference between the three GRMs we tested (1, panel B vs figure S6, panels A and B; table 1): all F yielded efficient (we use efficient to describe an estimate with low RMSE, thus which is unbiased and has low variance) estimates of b when used inside a LMM, except for F_{UNI}^u that slightly overestimates the strength of ID while F_{PED} slightly underestimates it. This suggests that large sample sizes (here 11,924 individuals) combined with a mixed model allow efficient ID estimation regardless of the F used. The three mixed models, however, perform less efficiently when the sample size is reduced, as we demonstrate with both subsampled PEDIGREE populations ($n = 2,500$): many replicates produced estimates above zero for b (figure 1, panels D and F; figure S6, panels C to F; table 1). RMSE were particularly large for F_{PED} , $F_{HBD.100KB}$ and $F_{ROH.100KB}$ with the mixed model using the unweighted GCTA GRM matrix (LMM_{GCTA}^u) (figure S6, panel D; table 1) but that is because most subsampled individuals F_{PED} , $F_{HBD.100KB}$ and $F_{ROH.100KB}$ were close to zero. Additionally, increasing the variance of subsampled

individuals' F (i.e. ranged subsampling) led to better estimates of b with reduced variance among replicates compared to random subsampling (figure 1, panels D vs F: figure S6, panels C vs E and D vs F, table 1). In order to assess the performance of the different models with even smaller sample sizes, which are common for wild and non-model species, we simulated pedigrees with 50, 100, 250, and 500 individuals (figure S7). With all sample sizes, the simple LM produces biased estimates (figure S7, panels A, E, I, and M). Including a GRM improved the estimation of b , but not as well as it did for large sample sizes (figure S7, panels B, C, D, F, G, H, J, K, L, N, O and P). The lowest RMSE was obtained with LMM_{AS} , but the difference with both $GCTA$ -based GRMs was marginal.

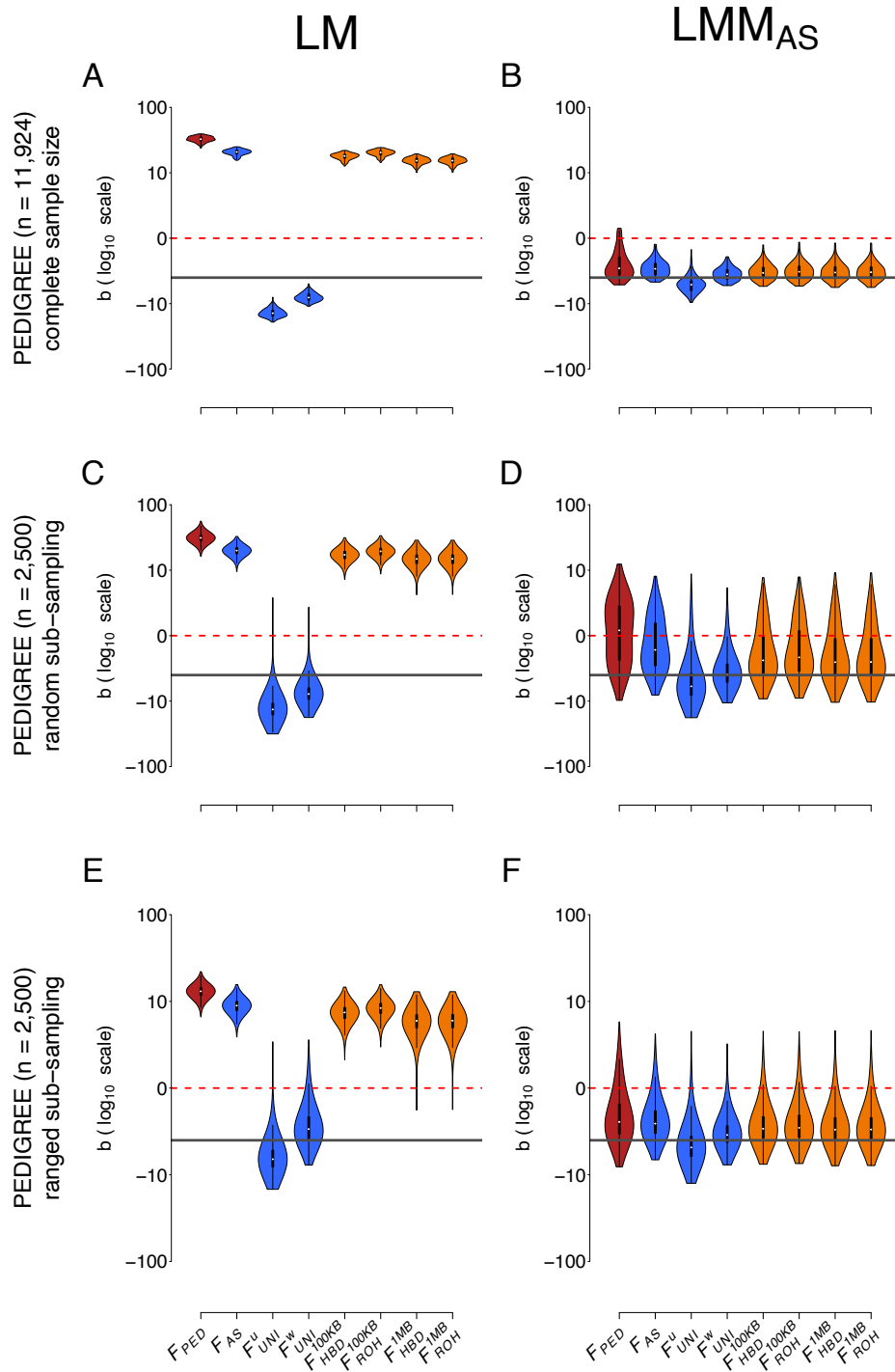


Figure 1: Comparison of the estimation of inbreeding depression strength (b) among different F estimates and two models in the PEDIGREE population. Each column represents a regression model. The first column depicts the simple linear regression (LM) (panels A, C and E) and the second column depicts the linear mixed model with the allele sharing relatedness matrix as a random component (LMM_{AS}) (panels B, D and F). The first row represents the complete simulated population (11,924 individuals, panels A and B). The second row shows the random subsampling (2,500 individuals, panels C and D). The third row shows the ranged subsampling (2,500 individuals, panels E and F). Inbreeding estimates presented in this graph are F_{PED} , F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$ and finally $F_{ROH.1MB}$. For panels A and B, violin plots show the distribution of the inbreeding depression strength estimates (b) among the 100 simulation replicates. For panels C to F, violin plots represent the distribution of the inbreeding depression

strength estimates (b) for the 10,000 simulation and subsampling replicates (100 subsampling replicates for each of the 100 simulation replicates). The solid dark grey line is the true strength of ID ($b = -3$). The dashed red line represents the absence of ID ($b = 0$), meaning that we failed to detect ID in any replicate above this line. Note that all panels are in \log_{10} scale.

Table 1: RMSE on \mathbf{b} estimate in the PEDIGREE population. These values are for the complete ADD & DOM & DEMA scenario. See tables S3 – S6 for the other scenarios.

Model	PEDIGREE	F _{PED}	F _{AS}	F_{UNI}^u	F_{UNI}^w	F _{HBD.100KB}	F _{ROH.100KB}	F _{HBD.1MB}	F _{ROH.1MB}
LM	Complete	34.82	22.71	10.17	4.17	19.93	22.22	17.4	17.4
LMM _{AS}	Complete	1.62	1.27	1.89	0.87	1.07	1.12	1.11	1.11
LMM _{GCTA^w}	Complete	1.62	1.27	1.89	0.87	1.07	1.12	1.11	1.11
LMM _{GCTA^u}	Complete	1.58	1.28	1.85	0.88	1.08	1.12	1.08	1.08
LM	Random sub	33.84	22.20	10.41	4.47	19.53	21.72	17.24	17.28
LMM _{AS}	Random sub	4.01	2.97	3.82	1.83	2.57	2.73	2.56	2.57
LMM _{GCTA^w}	Random sub	4.01	2.97	3.82	1.83	2.57	2.73	2.56	2.57
LMM _{GCTA^u}	Random sub	>1,000	2.75	3.44	1.78	> 1,000	> 1,000	> 1,000	> 1,000
LM	Ranged sub	15.22	11.04	3.46	1.61	9.58	10.52	8.13	8.15
LMM _{AS}	Ranged sub	2.09	1.82	2.13	1.26	1.61	1.67	1.58	1.58
LMM _{GCTA^w}	Ranged sub	2.09	1.82	2.13	1.26	1.61	1.67	1.58	1.58
LMM _{GCTA^u}	Ranged sub	>1,000	1.69	2.05	1.24	> 1,000	> 1,000	1.53	1.54

1,000 Genomes Project (1KG)

Figure 2 illustrates the estimates of ID strength (b) for the different inbreeding coefficients (F), when using either a LM or a LMM for two subsets of the 1,000 Genomes Project: EAS and AFR, as well as for the entire world population. It has the same structure as figure 1. Root mean square error (RMSE) values associated with both models and populations can be found in table 2. Interestingly, we see little difference between LM and LMM and the different GRMs when there is no structure among the samples even with small sample sizes (EAS: figure 2, panel A and B vs figure S8, panels A and B; table 2; AFR: figure 2, panel C and D vs figure S8, panels C and D; table 2). Similarly to what was observed for the PEDIGREE population, when some structure exists (population structure in the WORLD population compared to family structure in the PEDIGREE population), the simple LM fails to accurately estimate the strength of ID, regardless of the F (figure 2, panel E; table 2). In contrast to the pedigree population showing no difference between the three GRMs (figure 1 and figure S6), the most efficient estimates of b are obtained only with the LMM_{AS} model and with F_{UNI}^w in the highly structured WORLD population (figure 2, panel F vs figure S8 panels E and F; table 2). In fact, the models including the $GCTA^w$ and $GCTA^u$ matrices cannot efficiently estimate b with any of the inbreeding coefficients: even though b with F_{UNI}^w are unbiased, the variance is very large (panel F; figure S8, table 2). In addition, several replicates did not converge when both $GCTA^w$ and $GCTA^u$ models were used which was never the case

with the GRM_{AS} . Numbers of such replicates are indicated in the figures' legend and in supplementary tables S6 – S8. Similarly to what was done for the PEDIGREE population, we subsampled individuals from the WORLD population to test the different models with smaller sample sizes (50, 100, 250 and 500, as shown in figure S9). The results are very similar to those observed in the large WORLD population. Unsurprisingly, the simple LM fails to adequately quantify inbreeding depression with all sample sizes (figure S9 panels A, E, I and M), and the most effective estimation of b is obtained using LMM_{AS} and F_{UNI}^w . (figure S9 panels C, G, K and O). Both GCTA-based models completely fail to accurately quantify b with any F (figure S9 panels C, D, G, H, K, L, O and P).

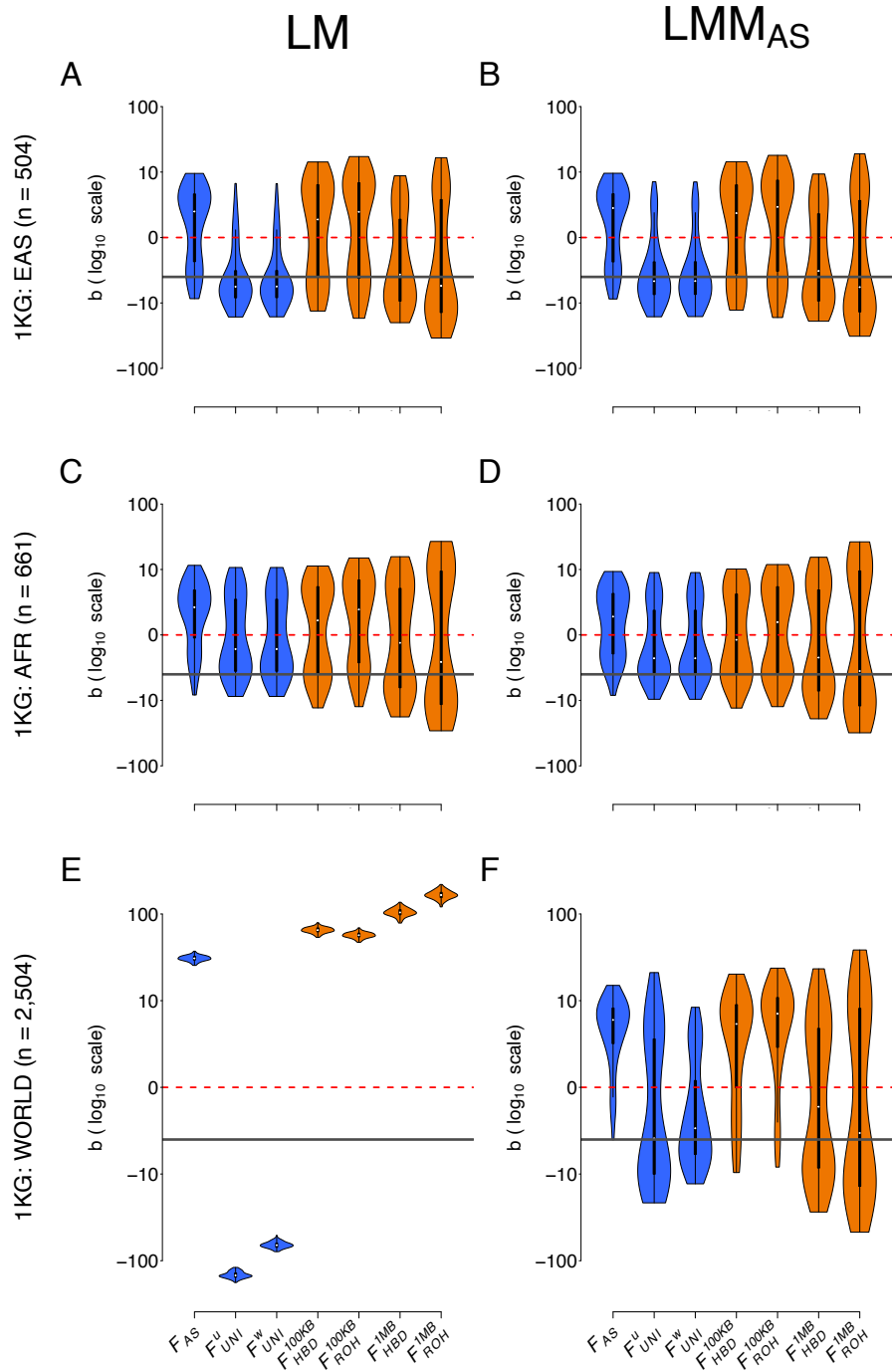


Figure 2: Comparison of the estimation of inbreeding depression strength (b) among different F estimates and two models in the three populations from the 1,000 Genomes project. Each column represents a regression model. The first column depicts the simple linear regression (LM) (panels A, C and E) and the second column depicts the linear mixed model with the allele-sharing relatedness matrix as a random component (LMM_{AS}) (panels B, D and F). The three rows correspond to the three populations from the 1,000 Genomes project: EAS on panels A and B, AFR on panels C and D and WORLD on panels E and F. Inbreeding estimates presented in this graph are F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$ and finally $F_{ROH.1MB}$. Violin plots show the distribution of the inbreeding depression strength estimates (b) among the simulation 100 replicates. The solid dark grey line is the true strength of ID ($b = -3$). The dashed red line represents the absence of ID ($b = 0$), meaning that we failed to detect ID in any replicate above this line. Note that all panels are in \log_{10} scale.

Table 2: RMSE on b estimate in the three 1,000 Genomes Project populations: EAS, AFR and WORLD. These values are for the complete ADD & DOM & DEMA scenario. See tables S3 – S6 for other scenarios.

Model	1 KG	F_{AS}	F_{UNI}^u	F_{UNI}^w	FHBD.100KB	FROH.100KB	FHBD.1MB	FROH.1MB
LM	EAS	5.55	4.9	4.86	7.14	7.93	6.19	10.58
LMM _{AS}	EAS	5.67	4.68	4.64	7.41	8.22	6.12	10.39
LMM _{GCTA} ^W	EAS	5.67	4.68	4.64	7.28	8.06	6.11	10.39
LMM _{GCTA} ^U	EAS	5.48	4.74	4.71	7.1	7.87	6.18	10.57
LM	AFR	5.93	4.81	4.81	6.03	7.21	17.21	13.12
LMM _{AS}	AFR	5.15	4.07	4.07	5.46	6.2	7.15	13.1
LMM _{GCTA} ^W	AFR	5.15	4.07	4.07	> 1,000	> 1,000	7.16	13.1
LMM _{GCTA} ^U	AFR	5.78	4.42	4.42	5.92	6.93	7.2	13.11
LM	WORLD	32.91	142.95	62.21	67.42	59.15	107.67	169.73
LMM _{AS}	WORLD	8.63	8.34	4.17	9.15	10.97	8.78	14.6
LMM _{GCTA} ^W	WORLD	9.84	> 1,000	> 1,000	11.19	13.92	> 1,000	> 1,000
LMM _{GCTA} ^U	WORLD	18.18	> 1,000	> 1,000	27.52	26.91	> 1,000	> 1,000

Comparing inbreeding coefficients

With both the LM and LMMAS models in the three populations from the 1,000 Genomes Project (EAS, AFR and WORLD, panels A – F) and for the LM in the PEDIGREE population, F_{AS} is consistently underestimating the strength of ID, particularly when there is strong structure (WORLD: figure 2, panels E and F). It is because DEMA is included in the model and strongly influences the quantification of ID by F_{AS} . In the absence of a DEMA, F_{AS} produces efficient estimates (figures S10 – S13). In addition, F_{AS} is sensitive to the dominance effects being proportional to MAF but to a lesser extent and in the opposite direction (Figure S10 vs. Figure S11). Concerning the other SNP-based F , F_{UNI}^u is constantly overestimating the strength of ID and is the most sensitive to population structure: its variance is much larger compared to F_{UNI}^w in the structured WORLD population and with all models (figure 2, panel F; table 2). Interestingly, the variance of F_{UNI}^u is affected only when allele effect sizes and/or dominance coefficients are proportional to MAF, but not by DEMA (Figures S10 – S17). In contrast, F_{UNI}^w is the least sensitive to allele effect sizes or dominance coefficients proportional to MAF and DEMA (figures S10 – S17), which makes it the most appropriate F for estimating ID (figure 2, panel F; table 2). Since the difference between F_{UNI}^u and F_{UNI}^w is the weight given to rare and common alleles, we conducted the same analyses (including the re-estimation of both F and GRMs estimation) on the WORLD population but excluding loci with $MAF > 0.05$ and showed that there is no difference between both F_{UNI} when rare alleles are removed (figure S18). Concerning the F calculated from ROHs and HBD segments, there is not much difference between PLINK and BCFTOOLS except for the variance among b estimates, which is slightly smaller with BCFTOOLS compared to

PLINK (figure 2, panels A – F; table 2). In addition, focusing on recent inbreeding by including only large segments (here larger than 1MB) yielded better results in the WORLD population (figure 2, panel F). Since BCFTools is a model-based approach, there is no mandatory length requirement for HBD segments. In light of this, we also estimated F_{HBD} based on HBD segments without any size restrictions, and the results are similar to those obtained using $F_{\text{HBD},100\text{KB}}$ (figure S19). We also quantified inbreeding depression with ROHs and HBD segments larger than 5MB but it did not improve the estimation of b (figure S19).

Comparing genetic relatedness matrices

Since we identified F_{UNI}^W as the best inbreeding coefficient to quantify ID, figure 3 contrasts the four different models for this coefficient in the four populations: each panel corresponds to one population. As mentioned above, there is almost no difference among the different GRM matrices in the extremely large complete PEDIGREE population (figure 3, panel A; table 1) and between any of the models in the two homogeneous populations (EAS and AFR) (figure 3, panels B and C; table 2). However, in the highly structured WORLD population, LMM_{AS} gives the most efficient result due to its smaller variance and RMSE (figure 3, panel D; table 2).

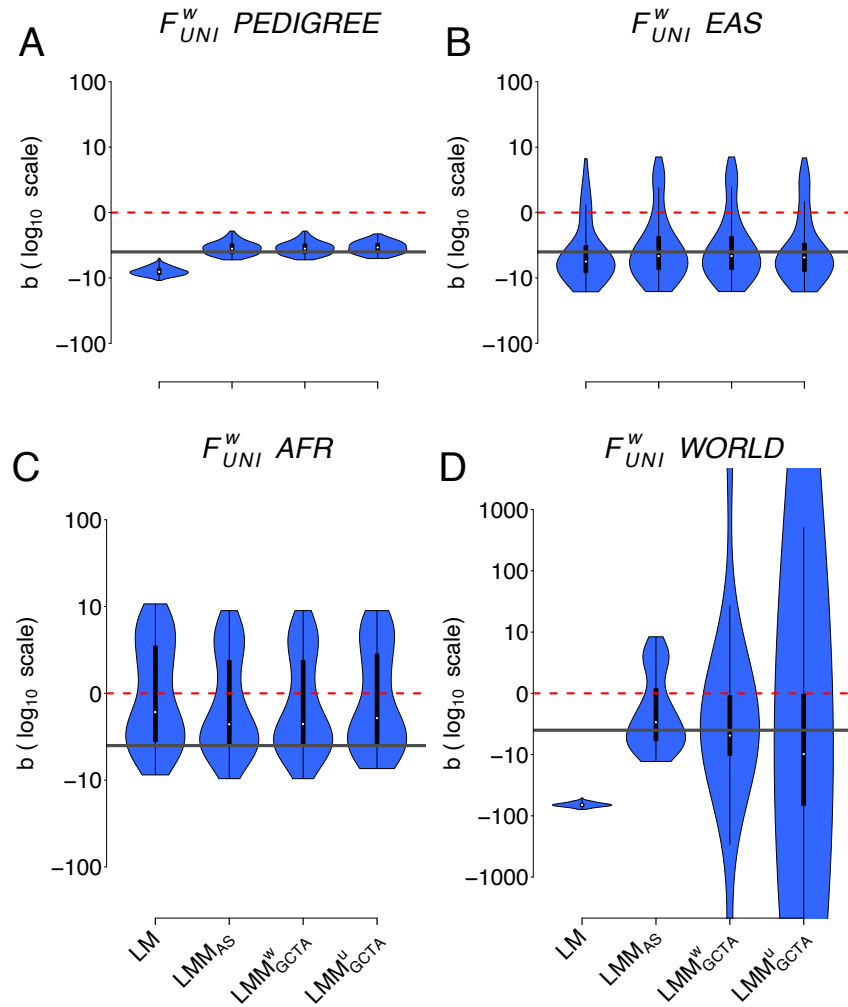


Figure 3: Comparison of the inbreeding depression strength estimates (b) with F_{UNI}^w in the four populations with four different models. The four models are: i) the simple linear regression (LM), ii) the linear mixed model with the allele sharing relatedness matrix as a random factor (LMM_{AS}), iii) the linear mixed model with the weighted GCTA relatedness matrix as a random factor (LMM_{GCTA}^w), and iv) the linear mixed model with the unweighted GCTA relatedness matrix as a random factor (LMM_{GCTA}^u). Panel A shows the simulated PEDIGREE population, panel B the EAS population, panel C the AFR population and finally panel D the WORLD population. Note that all panels are in \log_{10} scale. Also note that LMM did not converge for some replicates (yielding estimated b values above 1,000 or below -1,000, not shown in the graph). Percentages of replicates which did not converge: panel D (WORLD): 21% for GMR_{GCTA}^w 20% for GRM_{GCTA}^u.

Distribution of additive and dominance effects

We found a difference between the three linear mixed models only because the scenario presented in the main text includes effect sizes and dominance coefficients proportional to causal markers' MAF as well as DEMA. When none of these three parameters are included, there is little difference between the three linear mixed models (figure S10, panels B, F, J, N vs panels C, G, K, O vs panels D, H, L, P; tables S3 – S6). Additional simulations were conducted without additive and dominance coefficients proportional to

loci's MAF and DEMA to assess their impact on ID detection. The individual and pairwise effects of additive and dominance coefficients being proportional to MAF and DEMA (the other scenarios of table S1) are explored and discussed in details in supplementary material and figures S10 – S17.

Finally, we also investigated i) the effect of the LDMS stratification method proposed by Yengo *et al.* [50] (figures S10 – S17) but found that it improves results only with the simple LM and ii) the effect of using intermediate frequencies causal loci (figure S20) which reduced the variance in b estimates for all inbreeding coefficients.

Application to empirical data

As an illustration of our method, we analyzed adult mass and bill depth of a metapopulation of house sparrows in northern Norway using a dataset from Niskanen *et al.* [38] (analyses for other morphological traits are given in supplementary material). For mass (table 3), the slope associated with F_{UNI}^W is $b = -2.39$ and its associated p-value = 0.02 in the simple LM. The models the authors of the paper used [38] is a linear mixed model with the island and year nested in islands as random effects and results in $b = -1.98$ and a associated p-value = 0.05. Using only GRM_{AS} as a random effect makes the slope steeper and more significant: $b = -2.86$, p-value = 0.007. If we include the GRM_{AS} , the island and year nested in island (the full model), the results are very similar: $b = -2.85$, p-value = 0.006. For bill depth, the slope associated with F_{UNI}^W is positive (table 3) and significant for the LM ($b = 0.27$, p-value = 0.039), which suggests the presence of outbreeding depression for this trait. With the LMM_{AS} , however, the slope is shallower and not significant ($b = 0.22$, p-value = 0.106). Including islands and years (nested in islands) as random effects show a similar pattern and the full model makes the slope for F_{UNI}^W shallower and its p-value larger.

Table 3: Analysis of adult mass and bill depth from 1,786 adult sparrows. LM: simple linear model with Sex and F_{UNI}^w as explanatory variables. LMM_{AS}: linear mixed model with sex and F_{UNI}^w as fixed effects and GRM_{AS} as random effect. LMM: linear mixed model with sex and F_{UNI}^w as fixed effects and island and year nested in island as random effects. LMM_{FULL}: linear mixed model with sex and F_{UNI}^w as fixed effects and island, year nested in island and GRM_{AS} as random effects. V_I : variance component of island effect; $V_{Y:I}$: variance component for year nested in island; V_A : additive variance; V_E : residual variance; P_F : p-value for the slope b of F_{UNI}^w to be 0.

Mass	Int.	Sex	F	V_I	$V_{Y:I}$	V_A	V_E	P_F
LM	33.0	-1.39	-2.39				4.59	0.02
LMM _{AS}	34.3	-1.41	-2.86			1.56	3.02	0.007
LMM	32.9	-1.38	-1.98	0.15	0.27		4.27	0.050
LMM _{FULL}	34.3	-1.40	-2.85	0.01	0.17	1.45	2.92	0.006

Bill depth	Int.	Sex	F	V_I	$V_{Y:I}$	V_A	V_E	P_F
LM	8.1	0.04	0.27				0.08	0.039
LMM _{AS}	8.1	0.03	0.22			0.04	0.04	0.106
LMM	8.1	0.04	0.24	0.00	0.01		0.07	0.068
LMM _{FULL}	8.1	0.03	0.23	0.00	0.01	0.04	0.04	0.084

Discussion

By analyzing the phenotypes of a large simulated pedigreed polygamous population with strong family structure as well as subsets of the 1,000 genomes project [44], we demonstrated that, despite population or family structure, inbreeding depression estimates can be efficiently estimated if the data are analyzed with a mixed model including the genomic relationships among individuals as a random effect. While the use of a relationship matrix as a random factor in mixed models for quantitative genetics analyses is standard [30], and genomic relationship matrices (GRMs) have been used for the estimation of heritability [48, 25, 49, 19] and in GWAS [9, 48, 29, 26, 16, 19] for a long time, the application of LMM including GRMs to quantify inbreeding depression is seldom used (In 2012, McQuillan and collaborators [32] employed the `GenABEL` R package developed for GWAS in order to run inbreeding depression models including an IBD-based kinship matrix; we did not discover any follow-up papers using a similar approach until Nishio *et al.* [37] who used the `GCTA`-based GRM in 2023, although Stoffel *et al.* [41] used a model with breeding values as random effects). To our knowledge, we are the first to evaluate the ability of the LMM approach (including different GRMs) to quantify ID and compare it to the classical LM. First, we show that for most scenarios, ID is better estimated with LMM than with a simple LM and secondly, compared to other GRMs in LMM, the allele-sharing based GRM provides the most efficient results, especially for small sample sizes and samples with a high family or

population structure. In addition, among the several inbreeding estimators tested F_{UNI}^w proved to be the most reliable coefficient to quantify inbreeding depression.

We observed trivial differences among the different models when there is no population structure (i.e. in the EAS and AFR populations). However, as soon as there is some structure (the WORLD and PEDIGREE populations) the classical linear model (LM) completely fails to estimate b regardless of the inbreeding coefficient used. This result is concordant with Yengo *et al.* (2017) [50] where the authors quantified ID using a simple linear model and demonstrated that F_{HOM} (whose properties are very similar to F_{AS}), F_{UNI}^w and two different F_{ROH} were sensitive to population structure. As for the comparison of three linear mixed models (LMM), they perform equally when the population structure is weak (EAS and AFR) and with very large sample sizes (11,924 individuals from the complete PEDIGREE population). Although samples of this size are common for research on humans, they will seldom be found in wild populations. We therefore subsampled the PEDIGREE population to 2,500 individuals in order to investigate the effect of a smaller sample size and the range of inbreeding of the samples. We used two types of subsampling: i) random subsampling where individuals were chosen completely randomly and ii) ranged subsampling where individuals were chosen to maximise the range of F in the sampled population. As expected, when we subsampled individuals from the PEDIGREE population, RMSE values associated with b estimation increased slightly for both LMM_{AS} and LMM_{GCTA}^w mixed models and we failed to detect ID in some replicates. Accordingly, even with 2,500 individuals, we lack power and several thousands of individuals would be required to detect ID efficiently as Keller *et al.* and Caballero *et al.* previously pointed out [23, 5]. With the LMM_{GCTA}^u mixed model, all inbreeding coefficients but F_{AS} and F_{UNI} had convergence issues, suggesting that the LMM_{GCTA}^u mixed model is the least robust of the three mixed models. As expected, randomly subsampling individuals lead to a larger variance of b estimates compared to the ranged subsampling scheme, indicating that maximizing the variance of samples' F improves the estimation of b , although it is not obvious how such sampling could be done in non-monitored natural populations. It is unlikely that the sample sizes we used (2,500 individuals) will be found in many wild species, particularly in endangered populations where monitoring inbreeding and inbreeding depression is critical. Therefore, we re-simulated small pedigrees (of 50, 100, 250, and 500 individuals). Interestingly, all three LMM models were almost as efficient with small sample sizes as they were with pedigrees that were subsampled but the simple LM completely failed to accurately

estimate b . In part, this can be attributed to the relatively weak familial genetic structure (the PEDIGREE population has relatively weak family structure compared to the WORLD strong population structure). Additionally, the sampled individuals are highly inbred which makes the detection of inbreeding depression easier.

When we add a strong population structure in addition to small sample size (2,504, 500, 250, 100 and 50 individuals) from the highly structured WORLD population), we observe striking differences between the three different GRMs. The linear mixed model including the allele-sharing based GRM (LMM_{AS}) resulted in the most efficient estimations of b . In addition, the mixed models with both GRM_{GCTA}^u and GRM_{GCTA}^w did not converge for a high percentage of replicates (compared to 0% for LMM_{AS}) emphasizing that LMM_{AS} is the best model for quantifying inbreeding depression in highly structured populations and that it can also be applied to small sample sizes. This is because the allele-sharing based GRM matrix is a better estimator of kinship compared to both GCTA matrices [14, 46]. Indeed, what the GRM_{AS} estimates is the actual kinship in the population, based on how many alleles individuals share. In contrast, what both GRM_{GCTA}^w and GRM_{GCTA}^u estimate is a combination of individual kinship, their mean kinship with the other individuals and the overall mean kinship in the population (see equation 3 in Goudet *et al.* [46]). Consequently, since the kinship itself is better estimated with GRM_{AS}, the non-independence of observations (and thus the population structure) is better accounted for with LMM_{AS} which leads to better b estimates. Importantly, the inclusion of a GRM in the ID estimation model is not limited to simple linear models. Even though we used only linear models in this study, any type of generalized linear model can incorporate a GRM as a random factor. Consequently, this method can be applied to any trait distribution. Furthermore, by including the GRM-based random factor, the non-independence of observations is better accounted for than by including the population as a random factor, and no prior knowledge of the population structure is required.

Comparing F

Concerning the different inbreeding coefficients, we found F_{UNI}^w to be the best F for quantifying ID. Indeed, F_{UNI}^w was the only coefficient we tested which was not sensitive to neither additive and dominance effect sizes being proportional to MAF nor DEMA resulting in the least biased estimation of b . On the contrary, we found that F_{UNI}^u was influenced by the dominance effect sizes being proportional to MAF and by population

structure. In F_{UNI}^u estimation, the rare alleles associated with large dominance effect sizes add noise in the estimation of b . Similarly, when there is population structure, rare alleles which have a strong influence on F_{UNI}^u are likely to be private alleles which will strongly bias population-specific allelic frequencies and eventually F_{UNI}^u estimation. Importantly, F_{UNI}^u performed as well as F_{UNI}^w when we filtered on $MAF > 0.05$ (for F and all GRMs estimation). This is because F_{UNI}^u uses the average of ratios, which results in loci with small MAF strongly influencing the outcome. When these rare loci are filtered out, the estimated F is no longer biased. This explains why Yengo *et al.* [50] found that F_{UNI}^u was the best F for quantifying inbreeding depression with a homogeneous subset of the UK biobank dataset: they filtered on $MAF > 0.05$ leading to F_{UNI}^u estimation not being influenced by rare alleles with strong additive and/or dominance effect sizes. Concerning F_{AS} , we found that it was very sensitive to DEMA. This result is also concordant with Yengo *et al.* [50] who found that F_{HOM} (with properties very similar to F_{AS}) was sensitive to DEMA. In this paper, the authors explain that this sensitivity is due to F_{HOM} (and thus F_{AS}) correlating strongly with minor allelic count which will create a spurious association with inbreeding depression in the presence of DEMA. However, F_{AS} resulted in the most efficient estimates of b when DEMA was not included in the model, suggesting that it is the best F to estimate inbreeding for neutral regions, as was argued by Zhang *et al.* [52]. Finally, we found that ROHs and HBD segments-based F , namely F_{ROH} and F_{HBD} , performed poorly: underestimating the strength of inbreeding depression (positive b) or displaying very large variance among replicates. This result is in contradiction with Kardos *et al.* [20, 21] and Nietlisbach *et al.* [36] who found that F_{ROH} and F_{HBD} were better at quantifying inbreeding depression compared to SNPs-independent based F . However, Alemu *et al.* [2] and Caballero *et al.* [6] showed the best F actually depends on the history of the population. Indeed, they showed that F_{ROH} and F_{HBD} and to a lesser extent F_{HOM} were better at quantifying homozygosity at loci with common alleles. On the contrary, F_{UNI}^u was better at quantifying homozygosity at rare alleles. Alemu *et al.* [2] and Caballero *et al.* [6] propose that, in populations with low effective sizes, selection is weaker and deleterious alleles may be able to reach intermediate frequencies as a result of drift. Therefore, both F_{ROH} and F_{HBD} (and F_{HOM} in their analyses) should perform better in such populations. In our study, the standard scenario (with no ADD, no DOM and no DEMA) mimics what happens in such small populations and we found that F_{ROH} , F_{HBD} and F_{AS} (which has similar properties to F_{HOM}) performed better than F_{UNI}^u (which is the F_{UNI} they tested) in the highly structured WORLD population and to a

lesser extent in the family structured PEDIGREE population. With homogeneous populations, we do not observe any difference between these inbreeding coefficients. Nevertheless, this is consistent with Alemu and collaborators' [2] results, as they used families which consequently create structure. On the other hand, in populations with a large effective size, selection maintains deleterious alleles at low frequencies which explains why Yengo *et al.* (2017) found that F_{UNI} was the best F with the large UK biobank dataset and this is consistent with what we have found with the ADD & DOM & DEMA scenario which mimics what happens in populations with large effective sizes.

Conclusion

In this paper, we showed that the more efficient method for estimating inbreeding depression is to use a mixed model with an allele-sharing-based relatedness matrix as a random component but F_{UNI}^W , as the inbreeding coefficient to predict inbreeding depression. The most commonly used GRM (GRM_{GCTA}^U) results in biased and highly variable estimates of b in structured populations. We stress that even if the results are greatly improved by using the allele-sharing GRM and F_{UNI}^W , the variance among replicates is still large and no inbreeding depression is detected in several replicates ($b \geq 0$) in (all sample sizes for the) highly structured WORLD population as well as in the small and slightly admixed AFR population. Therefore, detecting efficiently inbreeding depression of the magnitude commonly found and that we simulated requires very large sample sizes with several thousand individuals, particularly in structured populations. Unfortunately, this might be hardly feasible for wild and/or endangered populations.

References

- [1] 46th European Mathematical Genetics Meeting (EMGM) 2018, Cagliari, Italy, April 18-20, 2018: Abstracts. *Human Heredity*, 83(1):1–29, 04 2018.
- [2] S.W. Alemu, N.K. Kadri, C. Harland, P. Faux, C. Charlier, A. Caballero, and T. Druet. An evaluation of inbreeding measures using a whole-genome sequenced cattle pedigree. *Heredity*, 126(3):410–423, 2021.
- [3] B.L. Browning and S.R. Browning. Detecting Identity by Descent and Estimating Genotype Error Rates in Sequence Data. *The American Journal of Human Genetics*, 93(5):840–851, 2013.
- [4] C. Béréanos, P.A. Ellis, J.G. Pilkington, and J.M. Pemberton. Genomic analysis reveals depression due to both individual and maternal inbreeding in a free-living mammal population. *Molecular Ecology*, 25(13):3152–3168, 2016.
- [5] A. Caballero, A. Fernández, B. Villanueva, and M.A. Toro. A comparison of marker-based estimators of inbreeding and inbreeding depression. *Genetics Selection Evolution*, 54(1):82, 2022.
- [6] A. Caballero, B. Villanueva, and T. Druet. On the estimation of inbreeding depression using different measures of inbreeding from molecular markers. *Evolutionary Applications*, 14(2):416–428, 2021.

- [7] F.C. Ceballos, S. Hazelhurst, D.W. Clark, G. Agongo, G. Asiki, P.R. Boua, F.X. Gómez-Olivé, F. Mashinya, S. Norris, J.F. Wilson, and M. Ramsay. Autozygosity influences cardiometabolic disease-associated traits in the AWI-Gen sub-Saharan African study. *Nature Communications*, 11(1):5754, 2020.
- [8] C.C. Chang, C.C. Chow, L.C. Tellier, S. Vattikuti, S.M. Purcell, and J.J. Lee. Second-generation *PLINK*: rising to the challenge of larger and richer datasets. *GigaScience*, 4(1):7, 2015.
- [9] Y. Choi, E.M. Wijsman, and B.S. Weir. Case-control association testing in the presence of unknown relationships. *Genetic Epidemiology*, 33(8):668–678, 2009.
- [10] C. Dandine-Roulland and H. Perdry. The Use of the Linear Mixed Model in Human Genetics. *Human Heredity*, 80(4):196–206, 2015.
- [11] H.P. Doekes, P. Bijma, and J.J. Windig. How Depressing Is Inbreeding? A Meta-Analysis of 30 Years of Research on the Effects of Inbreeding in Livestock. *Genes*, 12(6):926, 2021.
- [12] T. Druet, and M. Gautier. A model-based approach to characterize individual inbreeding at both global and local genomic scales. *Molecular Ecology*, 26(20):5820–5841, 2017.
- [13] J. Goudet. *hierfstat*, a package for r to compute and test hierarchical f-statistics. *Molecular Ecology Notes*, 5(1):184–186, 2005.
- [14] J. Goudet, T. Kay, and B.S. Weir. How to estimate kinship. *Molecular Ecology*, 27(20):4121–4135, 2018.
- [15] J. Hidalgo, A. Cesarani, A. Garcia, P. Sumreddee, N. Larios, E. Mancin, J.G. García, R. Núñez, and R. Ramírez. Genetic Background and Inbreeding Depression in *Romosinuano* Cattle Breed in Mexico. *Animals*, 11(2):321, 2021.
- [16] G.E. Hoffman. Correcting for Population Structure and Kinship Using the Linear Mixed Model: Theory and Extensions. *PLoS ONE*, 8(10):e75707, 2013.
- [17] J.I. Hoffman, F. Simpson, P. David, J.M. Rijks, T. Kuiken, M.A.S. Thorne, R.C. Lacy, and K.K. Dasmahapatra. High-throughput sequencing reveals inbreeding depression in a natural population. *Proceedings of the National Academy of Sciences*, 111(10):3775–3780, 2014.
- [18] J. Huisman, L.E.B. Kruuk, P.A. Ellis, T. Clutton-Brock, and J.M. Pemberton. Inbreeding depression across the lifespan in a wild mammal population. *Proceedings of the National Academy of Sciences*, 113(13):3585–3590, 2016.
- [19] W. Jiang, X. Zhang, S. Li, S. Song, and H. Zhao. An unbiased kinship estimation method for genetic data analysis. *BMC Bioinformatics*, 23(1):525, 2022.
- [20] M. Kardos, G. Luikart, and F.W. Allendorf. Measuring individual inbreeding in the age of genomics: marker-based measures are better than pedigrees. *Heredity*, 115(1):63–72, 2015.
- [21] M. Kardos, P. Nietlisbach, and P.W. Hedrick. How should we compare different genomic estimates of the strength of inbreeding depression? *Proceedings of the National Academy of Sciences*, 115(11), 2018.
- [22] J. Kelleher, A.M. Etheridge, and G. McVean. Efficient Coalescent Simulation and Genealogical Analysis for Large Sample Sizes. *PLOS Computational Biology*, 12(5):e1004842, 2016.
- [23] L. Keller, and D.G. Waller. Inbreeding effects in wild populations. *Trends in Ecology & Evolution*, 17(5):230–241, 2002.
- [24] M.C. Keller, P.M. Visscher, and M.E. Goddard. Quantification of Inbreeding Due to Distant Ancestors and Its Detection Using Dense Single Nucleotide Polymorphism Data. *Genetics*, 189(1):237–249, 2011.
- [25] Y. Kim, Y. Lee, S. Lee, N.H. Kim, J. Lim, Y.J. Kim, J.H. Oh, H. Min, M. Lee, H-J. Seo, S-H. Lee, J. Sung, N.H. Cho, B-J. Kim, B-G. Han, R.C. Elston, S. Won, and J. Lee. On the Estimation of Heritability with Family-Based and Population-Based Samples. *BioMed Research International*, 2015:1–9, 2015.
- [26] A. Korte, B.J. Vilhjálmsson, V. Segura, A. Platt, Q. Long, and M. Nordborg. A mixed-model approach for genome-wide association studies of correlated traits in structured populations. *Nature Genetics*, 44(9):1066–1071, 2012.
- [27] E. Lavanchy, and J. Goudet. Effect of reduced genomic representation on using runs of homozygosity for inbreeding characterization. *Molecular Ecology Resources*, 23(4):787–802, 2023.

- [28] A.-L. Leutenegger, B. Prum, E. Génin, C. Verny, A. Lemainque, F. Clerget-Darpoux, and E.A. Thompson. Estimation of the Inbreeding Coefficient through Use of Genomic Data. *The American Journal of Human Genetics*, 73(3):516–523, 2003.
- [29] C. Lippert, J. Listgarten, Y. Liu, C.M. Kadie, R.I. Davidson, and D. Heckerman. *FaST* linear mixed models for genome-wide association studies. *Nature Methods*, 8(10):833–835, 2011.
- [30] M. Lynch and Bruce Walsh. *Genetics and Analysis of Quantitative Traits*. Sinauer, 1998.
- [31] K. Martikainen, A. Sironen, and P. Uimari. Estimation of intrachromosomal inbreeding depression on female fertility using runs of homozygosity in Finnish Ayrshire cattle. *Journal of Dairy Science*, 101(12):11097–11107, 2018.
- [32] R. McQuillan, N. Eklund, N. Pirastu, M. Kuningas, B.P. McEvoy, T. Esko, T. Corre, G. Davies, M. Kaakinen, L.-P. Lyytikäinen, K. Kristiansson, A.S. Havulinna, M. Gögele, V. Vitart, A. Tenesa, Y. Aulchenko, C. Hayward, A. Johansson, M. Boban, S. Ulivi, A. Robino, V. Boraska, W. Igl, S.H. Wild, L. Zgaga, N. Amin, E. Theodoratou, O. Polasek, G. Girotto, L.M. Lopez, C. Sala, J. Lahti, T. Laatikainen, I. Prokopenko, M. Kals, J. Viikari, J. Yang, A. Pouta, K. Estrada, A. Hofman, N. Freimer, N.G. Martin, M. Kähönen, L. Milani, M. Heliövaara, E. Vartiainen, K. Räikkönen, C. Masciullo, J.M. Starr, A.A. Hicks, L. Esposito, I. Kolcic, S.M. Farrington, B. Oostra, T. Zemunik, H. Campbell, M. Kirin, M. Pehlic, F. Faletta, D. Porteous, G. Pistis, E. Widen, V. Salomaa, S. Koskinen, K. Fischer, T. Lehtimäki, A. Heath, M.I. McCarthy, F. Rivadeneira, G.W. Montgomery, H. Tiemeier, A.-L. Hartikainen, P.A.F. Madden, P. d’Adamo, N.D. Hastie, U. Gyllensten, A.F. Wright, C.M. van Duijn, M. Dunlop, I. Rudan, P. Gasparini, P.P. Pramstaller, I.J. Deary, D. Toniolo, J.G. Eriksson, A. Jula, O.T. Raitakari, A. Metspalu, M. Perola, M.-R. Jarvelin, A. Uitterlinden, P.M. Visscher, J.F. Wilson, and on behalf of the ROHgen Consortium. Evidence of inbreeding depression on human height. *PLOS Genetics*, 8(7):1–14, 2012.
- [33] R. McQuillan, A.-L. Leutenegger, R. Abdel-Rahman, C.S. Franklin, M. Pericic, L. Barac-Lauc, N. Smolej-Narancic, B. Janicijevic, O. Polasek, A. Tenesa, A.K. MacLeod, S.M. Farrington, P. Rudan, C. Hayward, V. Vitart, I. Rudan, S.H. Wild, M.G. Dunlop, A.F. Wright, H. Campbell, and J.F. Wilson. Runs of Homozygosity in European Populations. *The American Journal of Human Genetics*, 83(3):359–372, 2008.
- [34] R. Meyermans, W. Gorssen, N. Buys, and S. Janssens. How to study runs of homozygosity using *PLINK*? A guide for analyzing medium density SNP data in livestock and pet species. *BMC Genomics*, 21(1):94, 2020.
- [35] V. Narasimhan, P. Danecek, A. Scally, Y. Xue, C. Tyler-Smith, and R. Durbin. *BCFtools/RoH*: a hidden Markov model approach for detecting autozygosity from next-generation sequencing data. *Bioinformatics*, 32(11):1749–1751, 2016.
- [36] P. Nietlisbach, S. Muff, J.M. Reid, M.C. Whitlock, and L.F. Keller. Nonequivalent lethal equivalents: Models and inbreeding metrics for unbiased estimation of inbreeding load. *Evolutionary Applications*, 12(2):266–279, 2019.
- [37] M. Nishio, K. Inoue, S. Ogawa, K. Ichinoseki, A. Arakawa, Y. Fukuzawa, T. Okamura, E. Kobayashi, M. Taniguchi, M. Oe, and K. Ishii. Comparing pedigree and genomic inbreeding coefficients, and inbreeding depression of reproductive traits in Japanese Black cattle. *BMC Genomics*, 24(1):376, 2023.
- [38] A.K. Niskanen, A.M. Billing, H. Holand, I.J. Hagen, Y.G. Araya-Ajoy, A. Husby, B. Rønning, A.M. Myhre, P.S. Ranke, T. Kvalnes, H. Pärn, T.H. Ringsby, S. Lien, B.-E. Sæther, S. Muff, and H. Jensen. Consistent scaling of inbreeding depression in space and time in a house sparrow metapopulation. *Proceedings of the National Academy of Sciences*, 117(25):14584–14592, 2020.
- [39] J.E. Pryce, M. Haile-Mariam, M.E. Goddard, and B.J. Hayes. Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle. *Genetics Selection Evolution*, 46(1):71, 2014.
- [40] S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M.A.R. Ferreira, D. Bender, J. Maller, P. Sklar, P.I.W. de Bakker, M.J. Daly, and P.C. Sham. *PLINK*: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, 81(3):559–575, 2007.

- [41] M.A. Stoffel, S.E. Johnston, J.G. Pilkington, and J.M. Pemberton. Genetic architecture and lifetime dynamics of inbreeding depression in a wild mammal. *Nature Communications*, 12(1):2972, 2021.
- [42] P. Sumreddee, S. Toghiani, E.H. Hay, A. Roberts, S.E. Agrrey, and R. Rekaya. Inbreeding depression in line 1 Hereford cattle population using pedigree and genomic information. *Journal of Animal Science*, 97(1):1–18, 2019.
- [43] N.A. Swinford, S.P. Prall, S. Gopalan, C.M. Williams, J. Sheehama, B.A. Scelza, and B.M. Henn. Increased homozygosity due to endogamy results in fitness consequences in a human population. *Proceedings of the National Academy of Sciences*, 120(43):e2309552120, 2023.
- [44] The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature*, 526(7571):68–74, 2015.
- [45] P.M. VanRaden. Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science*, 91(11):4414–4423, 2008.
- [46] B.S. Weir, and J. Goudet. A Unified Characterization of Population Structure and Relatedness. *Genetics*, 206(4):2085–2103, 2017.
- [47] S. Wright. Coefficients of Inbreeding and Relationship. *The American Naturalist*, 56(645):330–338, 1922.
- [48] J. Yang, S.H. Lee, M.E. Goddard, and P.M. Visscher. *GCTA*: A Tool for Genome-wide Complex Trait Analysis. *The American Journal of Human Genetics*, 88(1):76–82, 2011.
- [49] R.-C. Yang. Genome-wide estimation of heritability and its functional components for flowering, defense, ionomics, and developmental traits in a geographically diverse population of *Arabidopsis thaliana*. *Genome*, 60(7):572–580, 2017.
- [50] L. Yengo, Z. Zhu, N.R. Wray, B.S. Weir, J. Yang, M.R. Robinson, and P.M. Visscher. Detection and quantification of inbreeding depression for complex traits from SNP data. *Proceedings of the National Academy of Sciences*, 114(32):8602–8607, 2017.
- [51] C. Zhang, P. Wang, D. Tang, Z. Yang, F. Lu, J. Qi, N.R. Tawari, Y. Shang, C. Li, and S. Huang. The genetic basis of inbreeding depression in potato. *Nature Genetics*, 51(3):374–378, 2019.
- [52] Q.S. Zhang, J. Goudet, and B.S. Weir. Rank-invariant estimation of inbreeding coefficients. *Heredity*, 128(1):1–10, 2022.

Chapter III: Too big to purge: no evidence of purging in the islands of the European barn owl (*Tyto alba*)

Authors: Eléonore Lavanchy^{1,2}, Tristan Cumer^{1,2}, Alexandros Topaloudis^{1,2}, Anne-Lyse Ducrest¹, Céline Simon¹, Alexandre Roulin¹ and Jérôme Goudet^{1,2}

¹ Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland

² Swiss Institute of Bioinformatics, University of Lausanne, Lausanne, Switzerland

Chapter status: not submitted

Authors contribution: EL conceptualized the study. A-LD and CS produced whole genome resequencing libraries. EL, TC and AT generated genotypes data. EL performed all down-stream analyses and drafted the manuscript. EL and JG wrote and revised the manuscript.

Genomic data generation for this chapter is described in Annex I.

Too big to purge: no evidence of purging in the islands of the European barn owl (*Tyto alba*)

Abstract

An important aspect of assessing endangered levels and managing conservation is the study of inbreeding status, identifying its origin as well as assessing the mutation load in wild populations. In this study, we used 502 barn owls from continental and island populations across Europe. In addition to comparing inbreeding status, we determine whether inbreeding is due to non-random mating or high co-ancestry within the population. We show that islands have higher levels of inbreeding than continental populations, and that this is mainly due to small effective population sizes rather than recent consanguineous mating. We assess the probability that a region is autozygous along the genome and show that this probability decreases as the number of genes present in that region increases. Finally, we look for evidence of reduced selection efficiency and purging in island populations. Among island populations, we found an increase in numbers of both neutral and deleterious minor alleles, possibly as a result of drift and decreased selection efficiency but we found no evidence of purging.

Introduction

Mutations are the ultimate source of variability in nature, serving as the foundation for evolution and natural selection. However, only a handful of mutations are beneficial for the fitness of an organism, whereas the majority of mutations are neutral or deleterious [41, 39, 20, 34]. Despite the effects of natural selection, deleterious mutations may still persist within populations. There are two main reasons for this persistence: (i) stochastic segregation of alleles (called drift), which can result in partially deleterious mutations being transmitted by chance and persisting through generations, and (ii) the recessive nature of many deleterious mutations, which render them invisible to selection when present in a single copy [8, 28]. In natural populations, the persistence and expression of such deleterious effects can be highly detrimental, although the contribution of each of these processes is still mostly uncharacterised in the wild.

Inbreeding (caused by mating between related individuals) can occur from mating among closely related individuals (recent coalescence events), as well as from more ancient

coalescence events (even when the population practices random mating [40]), such as a historical population bottleneck, which reduces the effective population size [6]. In the case of recent coalescence events, individuals usually inherit significant portions of their genomes that are identical-by-descent (IBD), transmitted by the same common ancestor [61].

As a result of inbreeding, the homozygosity load increases, and deleterious recessive effects are expressed, which can eventually result in reduced fitness - referred to as inbreeding depression [5, 54, 35, 46, 48]. While inbreeding depression negatively impacts population fitness, it can paradoxically serve as a mechanism for “purging” deleterious mutations. On the one hand, selection may be able to eliminate these deleterious recessive alleles from the population through purging [28, 68, 58, 27]. The effectiveness of purging is determined by the dominance and deleterious effect associated with each allele: highly deleterious and recessive alleles are more easily purged [25]. On the other hand, while long-term inbreeding and bottlenecks can result in the purging of highly deleterious recessive alleles, the purging of mildly deleterious alleles is more challenging [25, 23, 42, 28, 15, 64]. Indeed, even though a reduction in population size can lead to purging, it will also reduce the strength of natural selection and increase the effect of genetic drift [10, 19, 37, 21]. Consequently, an increase in inbreeding can result in complex deleterious alleles’ distributions with a decrease of highly deleterious alleles and an increase of mildly and lowly deleterious alleles [25, 23, 42, 67, 19].

There is considerable interest in studying the IBD segments shared between individuals for various purposes, such as inferring demography and population structure [56, 33], as well as identifying deleterious alleles with causal implications [49]. It should be noted, however, that accurate IBD segments determination typically requires phased genetic data, a computationally intensive and challenging task, particularly for non-model organisms. In contrast, detecting IBD segments within individuals is a much simpler task that does not require phased data (since we are examining homozygous segments within an individual, we do not need to know which copy of the chromosome each allele came from). In 2008, McQuillan *et al.* proposed looking for long stretches of homozygosity in polymorphic regions of the genome [48] as proxy for IBD segments within individuals. While short homozygous segments are likely to be homozygous by chance, the probability that long such segments are identical purely by chance is low, thereby increasing the likelihood that they are IBD. McQuillan *et al.* introduced the term “Runs

of Homozygosity” (ROHs) to describe these long homozygous stretches and showed that they are an informative proxy for an individual’s inbreeding status [48]. The distribution and length of these ROHs can also provide insight into the historical demographic processes of a population [6]. Indeed, the length of an IBD segment is linked to the time of coalescence [63, 61]. As a result, populations where individuals carry long ROHs may indicate recent consanguineous mating, whereas populations with many short ROHs may suggest an older previous bottleneck event [48, 6].

A common method for identifying ROHs involves a windows-based approach, in which consecutive homozygous regions are identified. This method is implemented in `PLINK` [55, 7]. To improve IBD segment identification accuracy, model-based approaches employing hidden Markov models (HMMs) were developed, such as `BCFTOOLS` [50] and `RZOOROH` [2]. IBD segments identified with model-based approaches will be hereafter called Homozygous-by-Descent (HBD) segments [17, 2, 18]. ROHs (or HBD segments) have been extensively studied in the wild for inbreeding characterization [62, 27, 36], inbreeding depression estimation [38, 62], demographic history inference [51] and homozygosity mapping [62].

The barn owl (*Tyto alba*) is a widely distributed nocturnal raptor found throughout Europe both on the mainland and on islands. Europe’s ecological history has been marked by significant climatic fluctuations, most notably during the last glacial maximum (LGM) approximately 20,000 years ago [22, 30]. As a result, many species, including barn owls, were forced to migrate to warmer southern regions during this period [31, 32]. Previous research has identified three refugia in southern Europe: i) the Iberian Peninsula [1, 4, 11], ii) the Levant and Anatolia [1, 4, 11] and finally iii) Italy and Greece [11] from which Europe was re-colonized when the climate warmed again [44, 11]. Like most recolonization, the recolonization of Europe by the barn owl probably occurred with bottlenecks at the front of the colonization side, followed by population expansion. This scenario has been shown to lead to an increased number in deleterious mutations per individual and a reduction in selection efficiency due to the smaller effective population size both with simulated data [52] and with empirical data such as out-of-Africa human expansion [29, 47] and salmons postglacial recolonization [59]. However, since the barn owl recolonized populations are on average geographically close to refugia populations (everything is happening at the European scale at most) and as there is strong direct gene flow between most recolonized populations and refugia populations [11], the deleterious

enrichment is likely not as strong as what we observed for human out-of-Africa colonization of Europe.

Water bodies can act as barriers to barn owl dispersal and gene flow [4, 44, 45, 11]. Consequently, islands populations exhibit increased isolation [45, 44, 12] along with higher levels of inbreeding [45, 12] compared to continental populations. This reduced gene flow in conjunction with a smaller population size can also reduce the effectiveness of natural selection. Because of its unique demography, which includes a variety of population sizes, and well-studied recolonization history, the barn owl is well suited for studying the effect of effective population size, inbreeding, and purging in the wild.

In this study, we analyze 502 high-quality sequenced genomes of barn owls from 19 different populations throughout Europe. We characterize the inbreeding status and landscape of HBD segments by using HBD segment-based inbreeding coefficients and distributions (i.e. how many small and large segments an individual or population contains). Island populations and, to a lesser extent, populations outside of refugium have experienced the founder effect through colonization. As a result, we expect their effective population sizes to be smaller, resulting in a reduction in selection efficiency. Hence, we hypothesize that deleterious mutations are more prevalent within islands and in recolonized populations. In addition, we look for evidence of purging of highly deleterious alleles in these same populations.

Materials and Methods

All codes used in this chapter can be found on [GitHub](#).

Sampling, sequencing, and genotyping

We used 502 barn owls (*Tyto alba*) from 19 populations for this project (details in table S1 and Figure 1): 346 individuals from Switzerland (CH), 15 from Grand-Britain (GB), 12 from Ireland (IR), 11 from continental Greece (GR) and the Aegean islands (AE) each, 10 from Denmark (DK), Israel (IS), Portugal (PT), Italy (IT), Crete (CT), Cyprus (CY), East canary (EC) and West canary (WC) each, 6 from Georgia (GE), 5 from France (FR), Serbia (SB) and the Ionian islands (IO) each, 3 from Morocco (MA) and Corsica (CO) each (details on sample sex and location can be found in table S1). Figure 1 shows the 502 sample locations; samples from continental populations are shown in purple and

individuals from island populations in blue. We further split the continental populations into: refugium populations, defined as populations which were present during the last glacial maxima and recolonized populations. Based on previous work about recolonization of Europe after the last glacial maxima [1, 4, 11, 44], we separated refugium populations as PT, MA, IT, GR and IS and recolonized populations as CH, FR, DK and SB. Although the Mediterranean islands were suitable habitats during the last glacial maxima [44], we excluded all islands from the comparison for refugia and recolonized populations due to their small size which could bias the results.

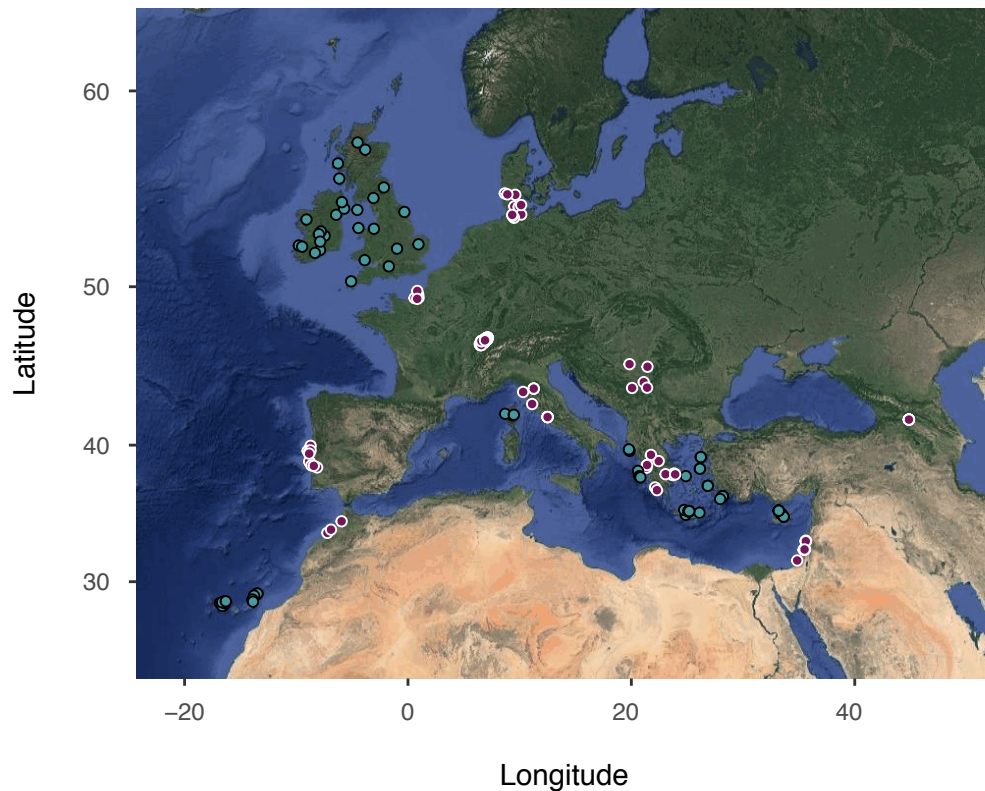


Figure 1: Map of samples' locations. Continental samples are shown in purple and island samples in blue.

For genomic data generation, we followed the same procedure as described in [44] and only used non-sexual chromosomes (details about the genotypic data generation are described in Annex I of this thesis). Briefly, we extracted DNA from blood and tissue samples with the DNeasy Blood & Tissue kit (Qiagen, Hilden, Germany). We then carried out whole-genome resequencing at the Lausanne Genomic Technologies Facility (GTF, University of Lausanne, Switzerland) using Illumina HiSeq 2500 PE high-throughput sequencing. Regarding the bioinformatics pipeline, we trimmed and aligned

raw reads to the reference genome (described in [44]) using `Trimomatic` (v0.39) [3] and `BWA-MEM` (v.0.7.17) [43]. We then performed ‘Base quality score recalibration’ (BQSR) with high-confidence calls from `GATK` and according to the procedure recommended for non-model species for which no set of “true variants” is available. We then called haplotypes with `GATK`’s (v4.2.6.1) *HaplotypeCaller* method and performed diploid joint-calling.

Variants were filtered following `GATK` ‘best practice’ procedure. *VariantFiltration* with `GATK` was used to apply the following filters to SNPs in the call set: $QD < 2.0$, $FS > 60.0$, $MQ < 40.0$, $MQRankSum < -12.5$, $ReadPosRankSum < -8.0$, $SOR > 3.0$. Afterwards, an additional ‘mappability’ filter was performed following: [these guidelines](#). To summarize, the rationale behind is to discard any regions for which the mapping quality was low. To this end, we split the reference genome into 150 bp regions (with one base-pair sliding-windows) and mapped these regions back to the reference genome where they came from. Regions for which the fraction of perfectly (and uniquely) mapping ‘reads’ was lower than 90% were then excluded with `VCFTOOLS` [13]. We then filtered individuals on genotypes’ depth (GD) with `BCFTOOLS` (v.1.15.1) [14]. We set all individuals’ genotypes with $GD < 5$ and $>$ mean individual depth plus three times it’s standard deviation to missing. We then filtered out any site with minor allele count (MAC) < 3 and fraction of missing individuals’ genotypes > 0.10 with `BCFTOOLS`. Finally, we only kept bi-allelic SNPs because it was needed for the R package we used for HBD segments characterization: `RZOOROH` [2]. We ended up with 14,093,173 bi-allelic high-quality SNPs.

HBD segments identification

HBD segments were called on autosomes using the `RZOOROH` package (v.0.3.1) [17, 2, 18]. Our model included 13 HBD classes and 1 non-HBD class with rates (R) of 2, 4, 8, 16, 32, 64, 128, 256, 502, 1024, 2048, 4096 and 8192 for the HBD classes and 8192 for the non-HBD class [17]. These HBD classes correspond to different coalescence event ages and the rate corresponds to the expected number of generations since the coalescence event divided by two. We considered the ‘most probable HBD segments’ estimated by `RZOOROH` Viterbi algorithm for HBD segments distributions analyses. HBD segment distributions were then obtained by quantifying the mean (among individuals) sum of HBD segment lengths falling into the autozygous HBD classes. We

considered a HBD class as autozygous if the rate was less than 1024 (i.e. if the coalescence event occurred during the last 512 generations).

To take the recombination rate along the genome into account, we used recombination maps built with `Lep-MAP3` [57] and described in Annex I of this thesis. We interpolated the genetic positions of our SNPs between each SNP present in the recombination map via a linear model. If no map was available for a specific super scaffold, we assumed a constant recombination rate of 2×10^{-8} (the average recombination rate between adjacent base pairs among the rest of the super scaffolds).

Estimating inbreeding coefficients

The F_{HBD} inbreeding coefficient is defined as the average probability (among markers) to belong to a HBD segment and were obtained with the `cumhbd` function from the `RZOOROH` package with a T value of 1024. This value means that only HBD segments coalescing less than 512 generations ago are considered autozygous (i.e. IBD).

F_{AS} is an allele-sharing-based estimator of inbreeding described in [66, 69]. It corresponds to the average allele-sharing for an individual (taking value 1 if the individual is homozygous and 0.5 if heterozygous at a specific locus), scaled by the mean allele-sharing between individuals of the population. Consequently, F_{AS} was estimated separately in each population. In addition, since the Swiss population (CH) contained related individuals, we only considered individuals with relatedness < 0.05 for estimating the mean between individuals allele-sharing for this population (the list of unrelated individuals can be extracted from table S1).

Probability of belonging to an HBD segment

We followed [62] to estimate the probability of belonging to an HBD segment along the genome. For this analysis, we excluded any super-scaffold with less than 10,000 SNPs. The probability to belong to an HBD segment was first estimated for each variant position as the sum of probabilities to belong to any HBD class estimated from the `hbdp` object from the output of the `zoorun` function from the `RZOOROH` R package. Average probabilities were then estimated via 20Kb overlapping sliding windows of 100Kb with the `windowscanr` R package.

Nucleotide diversity & Effective population size

The nucleotide diversity π was estimated with the *pi.dosage* function from the `hierfstat` R package. The effective population size (N_e) was then estimated as π divided by the mutation rate ($\mu = 4.6 \times 10^{-9}$ [60]). To obtain confidence intervals, we divided the genome in 1Mb segments and performed 1,000 bootstraps for each population separately.

Variant annotation

Variant annotation was performed with `SnPEff` [9] on the filtered complete dataset. Since *Tyto alba* `SnPEff` database was not available, we built the database first. We used the *build* method of `SnPEff` on the NCBI latest version of the *Tyto alba* assembly. Variant annotation was then performed with the *eff* method from `SnPEff` after filtering sites with missing data greater than 10%. Bi-allelic variants were classified into four categories using `SNPEff`: neutral, lowly deleterious, mildly deleterious, and highly deleterious. The neutral variants consist of changes to non-coding regions (including pseudo-genes), UTR regions or regions where it is difficult to predict the impact of the variant. Lowly deleterious variants refer to mutations that are harmless or unlikely to alter protein behavior, including synonymous mutations, nonsynonymous variants that change one or more amino acids but have similar properties to the originals, or changes in start (or stop) codons into different start (or stop) codon types. Mildly deleterious variants included mutations that may affect protein effectiveness, such as changes in the amino acid sequence leading to altered protein properties. Finally, highly deleterious mutations concerned variants that have a large (disruptive) impact on the protein, probably resulting in truncation or loss of function. The highly deleterious variants included, for example, loci implicated in protein-protein interactions (i.e. amino acids which are in contact within the same protein, possibly involved in structural conformation), rare amino acids that are likely to result in protein loss of function, variants mutating stop (or start) codons into non-stop (or non-start) codons (and vice versa). A more detailed description of these four categories can be found in the `SNPEff` documentation.

Accumulation of minor alleles

To examine whether our different populations are enriched in (expected homozygous) minor alleles, we used the R_{XY} and R^2_{XY} statistics described in [16].

These statistics aim at detecting an asymmetry in the number of minor alleles between two groups of genomes (X and Y) and count how many of these alleles (L_{XnotY}) or expected homozygous alleles (L^2_{XnotY}) are present in one group of genomes, but not the other and are formally defined as follow:

$$L_{XnotY} = \sum_i (d_X^i / n_X^i) (1 - d_Y^i / n_Y^i) \quad (1)$$

$$L^2_{XnotY} = \sum_i \frac{2d_X^i (n_X^i - d_X^i)}{n_X^i (n_X^i - d_X^i)} \left(1 - \frac{2d_Y^i (n_Y^i - d_Y^i)}{n_Y^i (n_Y^i - d_Y^i)} \right) \quad (2)$$

with d_X^i and n_X^i being respectively the number of (global) minor alleles and the total number of (haploid) genomes at site i in population X and d_Y^i and n_Y^i the number of minor alleles and the total number of (haploid) genomes at site i in population Y . The ratios between the two groups of populations are then calculated as:

$$R_{XY} = \frac{L_{XnotY}}{L_{YnotX}} \quad (3)$$

$$R^2_{XY} = \frac{L^2_{XnotY}}{L^2_{YnotX}} \quad (4)$$

We estimated both statistics for variants included in exons only and for each allele category (neutral, highly deleterious, moderately deleterious and lowly deleterious) separately. To account for different demographic histories and structure within each group of populations [16, 68, 27], we further divided these ratios by the same ratios estimated for intergenic variants only:

$$R'_{XY} = \frac{R_{XY}^{class}}{R_{XY}^{inter}} \quad (5)$$

$$R'^2_{XY} = \frac{R^2_{XY}^{class}}{R^2_{XY}^{inter}} \quad (6)$$

Consequently, the R'_{XY} and R'^2_{XY} statistics reported in the results section show an enrichment (if $R < 1$) or depletion (if $R > 1$) of (homozygous) minor alleles in population Y compared to population X in regards to what was observed for intergenic mutations. If natural selection has been as efficient in both populations, R'_{XY} and R'^2_{XY} should be equal to one. Similarly, if selection has been as efficient for lowly, mildly and highly deleterious alleles, the R'_{XY} and R'^2_{XY} ratios should be the same for each mutation category.

For this analysis, we compared the enrichment in minor allele rather than derived allele because we could not polarize our data. In order to avoid any bias due to sampling, minor alleles were estimated globally via 1,000 bootstraps by sampling only unrelated individuals. The allele which was identified as the minor allele in the majority of the bootstraps was used as the minor allele for this site. Polarization of a subset of 1,373,932 of the variants has been performed for another paper [44]. We used this subset to confirm whether the minor allele we identified corresponded to the inferred derived allele. Both methods identified the same allele for 87% of the sites. This fraction linearly decreased with minor allele frequency (being at 95% for sites with $MAF < 0.05$ and at 0.53 for sites with MAF between 0.45 and 0.5, illustrated in figure S1). As the majority of deleterious alleles are likely to be at low frequencies [53], we believe that the minor allele may be used as a proxy for the derived allele for this analysis.

Results

Inbreeding status

Individual inbreeding coefficients are shown in table S2. Figure 2 shows different metrics representing the level of inbreeding and coalescence time of parental relatedness in the different populations and population types (i.e. continental and island populations). The mean HBD segment based inbreeding coefficient was significantly higher for islands populations (mean $F_{HBD} = 0.101$) compared to continental populations (mean $F_{HBD} = 0.041$) (Wilcoxon rank test; $W = 3486$, $p\text{-value} < 2.2e-16$; large effect size: 0.525), (Figure 1, panel A). Interestingly the difference between continental and island populations starts showing with coalescence events older than 16 generations ago (corresponding to HBD class 4) (figure S2). Inbreeding coefficients distribution per population can be found in figure S3. There is little difference between the inbreeding coefficients for populations from continental refugium (mean $F_{HBD} = 0.041$; median F_{HBD}

= 0.039) and continental recolonized (mean $F_{HBD} = 0.041$; median $F_{HBD} = 0.032$, figure S4 panel A) but the distributions are still significantly different ($W = 10'063$, p -value = 0.01275; effect size: 0.122, considered small).

Panel B from Figure 2 contrasts individuals' HBD segments-based inbreeding coefficient: F_{HBD} to allele-sharing individual inbreeding coefficients: F_{AS} . F_{HBD} is the same as shown in panel A. F_{AS} represents the within individuals matching of alleles relative to the mean allele matching between individuals in the population. Fundamentally, this is a measure of how much alleles are associated at random in individuals (if F_{AS} is close to zero) rather than the same variant associated more often than expected by chance (if $F_{AS} > 0$). Individuals below the line for which F_{AS} is higher than F_{HBD} are subject to population structure (it indicates that the population is not homogeneous). On the contrary, individuals for which F_{HBD} is higher than F_{AS} are inbred but their inbreeding comes from ancient coalescence events (such as ancient population relatedness due to small effective population) rather than recent mating between relatives. In figure 2 panel B, we show that all individuals coming from island populations are above the identity line with an F_{HBD} higher than F_{AS} . This indicates that their inbreeding mostly comes from ancient relatedness likely due to small effective population size because of island isolation. In addition, among island populations, individuals from GB are closest to the identity line reflecting this island's larger size and thus higher population size. Within island populations, the most F_{HBD} inbred individuals come from CT, CY and EC. Four island individuals harbor low F_{HBD} and negative F_{AS} values (one from EC, one from CT, one from CY and finally one from AE). Finally, two island individuals (one from EC and one from GB) show very high F_{HBD} and F_{AS} suggesting that they come from mating between closely related individuals. Concerning continental populations, there is no strong difference between refugium populations and recolonized populations (except the Swiss individuals) (figure S4 panel B). Most Swiss individuals cluster below the identity line indicating the presence of unaccounted structure. Indeed, individuals in this population belong to families. If we get rid of family structure by trimming individuals with allele-sharing relatedness above 0.05, Swiss individuals are drawn closer the line (figure S5). Fourteen Swiss individuals are close to the identity line and truly inbred with high values of both F_{AS} and F_{HBD} . This is concordant with F_{PED} (described in SM) values obtained for the Swiss individuals (figure S6). PT individuals cluster closely together and close to the identity line. GE individuals cluster together with F_{HBD} between 0.05 and 0.1 and F_{AS} around -0.15. Concerning the

remaining continental populations, the three MA samples are below the identity line. We also observe few inbred individuals (Greeks, one Portuguese and one French) towards the top right part of the graph and the remaining individuals are non-inbred with F_{HBD} around 0.05 and F_{AS} around 0.

Panel C shows the mean number of HBD segments (N_{HBD}) according to the mean length of all HBD segments (S_{HBD}) for each individual. Inbred individuals (individuals with higher F_{HBD}) have more HBD segments but also longer segments. Island individuals have on average slightly more HBD segments compared to continental populations and longer HBD segments. Panel D shows HBD segments distribution for the island and continental populations. Very few long HBD segments (inbreeding events occurring in the last 8 generations) were found in either continental or island populations, suggesting that there is almost no recent inbreeding. However, island populations displayed a higher sum of lengths for medium-sized HBD segments coalescing between 16 and 128 generations (g) ago. Interestingly, refugium populations were slightly enriched in HBD segments coalescing 64g and 128g ago while recolonized populations were enriched in HBD segments coalescing 8g and 512g ago (figure S4 panel C and D). The increased sum of length of HBD segments coalescing 8g ago in the continental and refugia populations are driven by the few inbred Swiss individuals. F_{HBD} and HBD segments distribution estimations per population can be found in supplementary materials (figures S3 and S7).

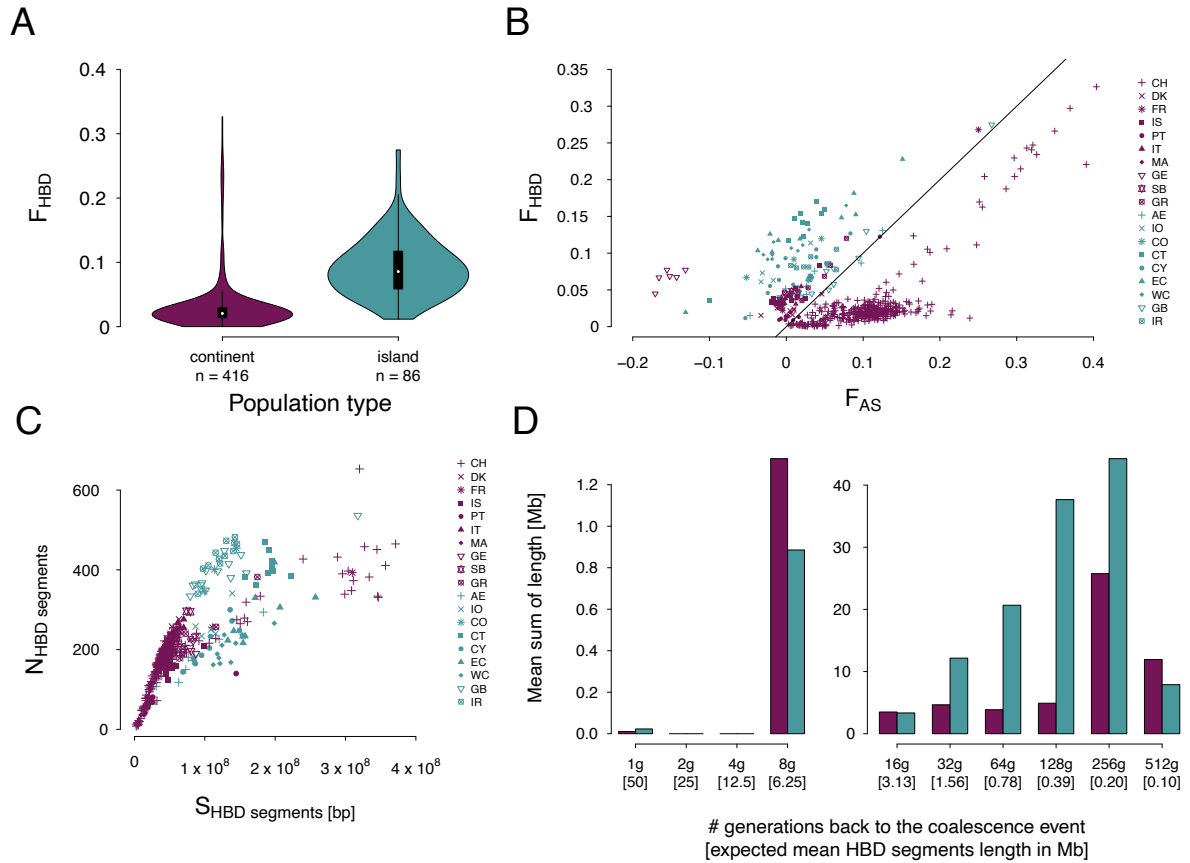


Figure 2: (For all panels, continental populations are shown in purple and island populations in blue) **A:** F_{HBD} distributions from continental and island populations. F_{HBD} considers a marker as autozygous if the coalescence event is up to 512 generations ago. **B:** scatter plot of F_{HBD} against F_{AS} . Each point represents one individual and its shape indicates which population it comes from. The black line is the identify line ($x = y$). **C:** number of HBD segments (N_{HBD}) as a function of the mean length of HBD segments (S_{HBD}) in base-pair. Each point represents one individual and its shape indicates which population it comes from. **D:** HBD segments distributions from continental populations and island populations. The y-axis represents the mean sum of length (among individuals) falling into the different categories of HBD segments (represented in the x-axis).

Effective population size

We also estimated the effective size N_e per population (table 1). The absolute values are very large but the relative comparisons should still be valid. On average, continental populations show higher N_e estimations compared to island populations (except the GE population, which is composed of related individuals, thus decreasing N_e estimation, the AE population, which has very strong gene flow with the GR population, and the CY populations). Within island populations, IR, followed by CT displayed the lowest N_e estimation. On the contrary, the AE and CY populations showed the highest N_e . Concerning continental populations, the refugium populations displayed the higher N_e estimation (except IT).

Table 1: Effective population sizes (N_e). N_e were estimated as the nucleotide diversity π divided by the mutation rate, estimated as 4.6×10^{-9} . We performed 1,000 bootstraps for all populations. Standard-error (SE) around the N_e mean indicates the variation among the different bootstraps.

Population Type	Population	Mean N_e	SE
Continent	PT	473,082	4,071
Continent	MA	466,349	3,911
Island	AE	458,560	3,901
Continent	IS	458,533	3,868
Continent	GR	457,446	3,873
Continent	CH	451,904	3,895
Continent	DK	451,414	3,934
Continent	IT	448,065	3,920
Island	CY	447,006	3,800
Continent	SB	446,922	3,891
Continent	FR	444,904	3,938
Island	WC	434,992	3,734
Island	IO	434,633	3,788
Island	GB	431,543	3,747
Island	EC	429,092	3,662
Island	CO	427,304	3,882
Island	CT	420,824	3,541
Island	IR	420,654	3,679
Continent	GE	367,102	3,388

Probability of belonging to an HBD segment

Figure 3 shows the probability of belonging to an HBD segment (of any class) along the genome (estimated as 100kb windows) as proposed in [62]. Probabilities are shown per super-scaffolds in panel A and the mean density function among all super-scaffolds is shown in panel B. For the vast majority of the genome, the probability to belong to an HBD segment is low (< 0.1). However, few super-scaffolds (such as super-scaffold 22) and specific genomic regions (such as the very beginning of super-scaffold 3) show particularly high HBD probabilities. We show in panel C that gene rich windows are very unlikely to be autozygous. Finally, we report regions with extremely high (HBD islands) and low (HBD deserts) HBD probabilities in tables S3 and S4. HBD islands were identified as windows in the top 2.5% of HBD probabilities. Similarly, HBD deserts were identified as the bottom 2.5% windows with lower HBD probabilities. Gene ontology enrichment analyses in HBD segments islands and deserts reveal no specific enrichment in these regions. We also investigated the genes present at the beginning of super-scaffold 3 (with the highest probabilities to be HBD) but could not identify any pattern.

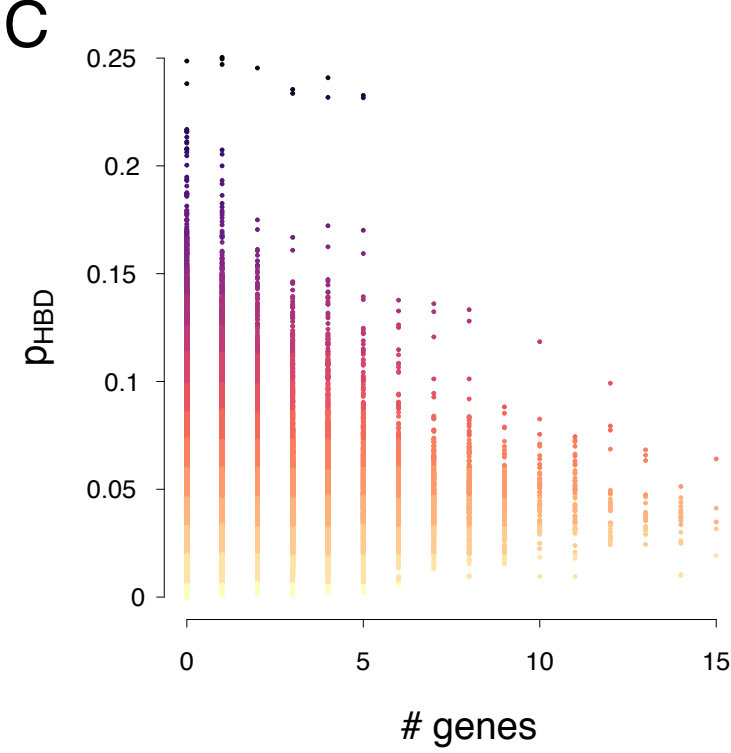
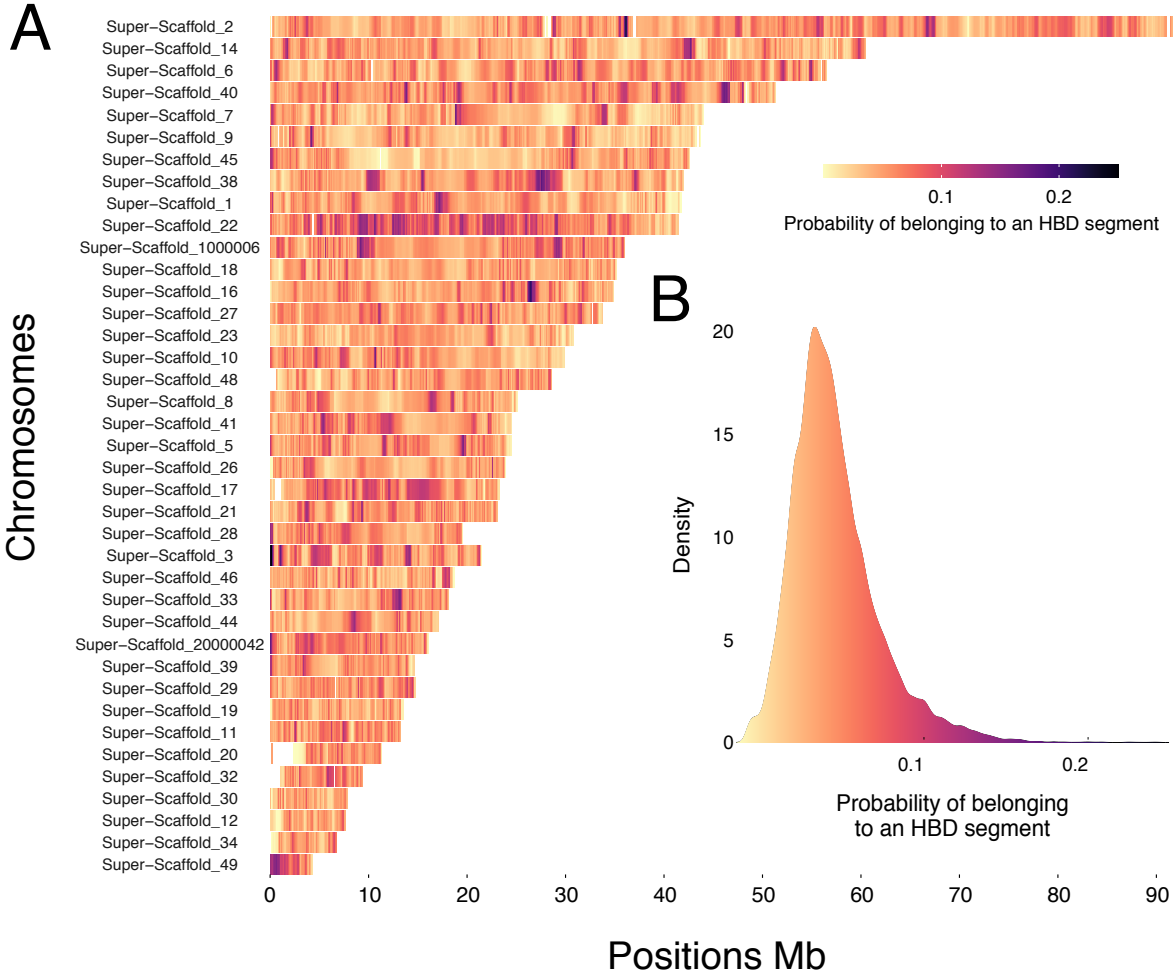


Figure 3: **A:** Probability to belong to an HBD segment coalescing less than 512 generations ago along the different super-scaffolds. Probabilities were estimated via overlapping sliding windows of 100 (+20) Kb. Blank spaces correspond to regions where no SNPs were present. **B:** Density of probability to belong to an HBD segment, all super-scaffolds included. The code to obtain this figure was obtained from [M. Stoffel GitHub](#). **C:** Probability that a 100Kb window is HBD according to the number of genes in this window.

Accumulation of minor alleles

Figure 4 panels A to D show the number of minor alleles per variants category (A: Neutral; B: Lowly deleterious, C: Moderately deleterious; D: Highly deleterious) in continental populations versus island populations. Island populations were significantly enriched in minor alleles for all variants categories (Wilcoxon rank sum tests; Neutral: $W = 6537.5$, $p\text{-value} < 2.2e\text{-}16$, effect size = 0.414, considered moderate; Lowly deleterious: $W = 7022.5$, $p\text{-value} < 2.2e\text{-}16$, effect size = 0.396, , considered moderate; Moderately deleterious: $W = 6907.5$, $p\text{-value} < 2.2e\text{-}16$, effect size = 0.400, considered moderate; Highly deleterious: $W = 8084$, $p\text{-value} = 1.189e\text{-}15$, effect size = 0.357, considered moderate). This enrichment was still present when we controlled for individual genetic diversity (by dividing the count of minor alleles by the number of polymorphic sites per individual) (figure S8, panels A to D). Interestingly, the lower tails of the islands violin plots always contained the individuals from GB and IR, suggesting that their minor allele counts are more similar to continental populations than to those from the other islands. Concerning the enrichment of each of these categories in respect to their enrichment in intergenic mutations, we consider a ratio significant if the standard errors around its estimation do not overlap with one. The R'_{XY} ratio (continents/islands) were not significantly different from one for all mutation types (panel E). This shows that island populations are similarly enriched in all allele categories compared to intergenic regions. Panels F to I show the number of homozygous minor alleles per variants category (F: Neutral; G: Lowly deleterious, H: Moderately deleterious; I: Highly deleterious) in continental populations versus island populations. Island populations were also significantly enriched in homozygous minor alleles for all variants categories (Wilcoxon rank sum tests; Neutral: $W = 4034.5$, $p\text{-value} < 2.2e\text{-}16$, effect size = 0.505, , considered large; Lowly deleterious: $W = 4482$, $p\text{-value} < 2.2e\text{-}16$, effect size = 0.489, , considered moderate; Moderately deleterious: $W = 4314.5$, $p\text{-value} < 2.2e\text{-}16$, effect size = 0.495, considered moderate; Highly deleterious: $W = 5269.5$, $p\text{-value} < 2.2e\text{-}16$, effect size = 0.460, considered moderate), even after correction for individual genetic diversity (figure

S8, panels E to H). The R'_{XY} ratios (continent/islands) were not significantly different from one for all homozygous minor alleles.

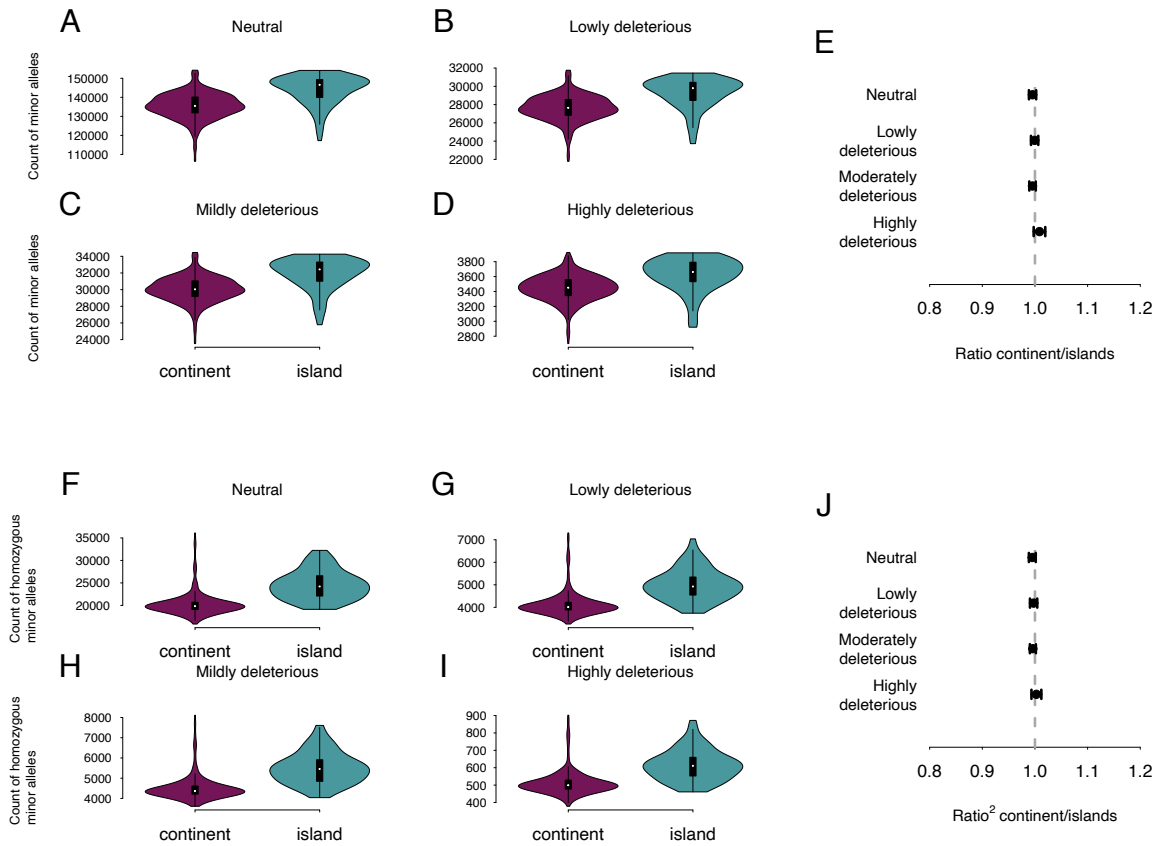


Figure 4: Distribution of minor alleles in continental versus islands populations. Minor alleles effects were classified with *SNPeff*. **A:** Count of neutral minor alleles. **B:** Count of lowly deleterious minor alleles. **C:** Count of moderately deleterious minor alleles. **D:** Count of highly deleterious minor. **E:** R'_{XY} ratio of minor alleles in continental populations compared to island populations scaled by the same ratio for SNPs located in intergenic regions. $R'_{XY} < 1$ indicates that islands populations are more enriched in minor alleles of the focal category compared to their enrichment for neutral intergenic alleles. On the contrary, $R'_{XY} > 1$ indicates that islands populations are depleted in minor alleles of the focal category compared to their enrichment for intergenic alleles. **F:** Count of homozygous neutral minor alleles. **G:** Count of homozygous lowly deleterious minor alleles. **H:** Count of homozygous moderately deleterious minor alleles. **I:** Count of homozygous highly deleterious minor alleles. **J:** R'^2_{XY} ratio of minor alleles in continental populations compared to island populations. $R'^2_{XY} < 1$ indicates that islands populations are more enriched in homozygous minor alleles of the focal category compared to their enrichment for neutral homozygous intergenic alleles. On the contrary, $R'^2_{XY} > 1$ indicates that islands populations are depleted in homozygous minor alleles of the focal category compared to their enrichment for homozygous intergenic alleles. For violin plots, continental populations are shown in purple and island populations in blue.

Similarly to figure 4, figure 5 panels A to D show the number of minor alleles per variants category (A: Neutral; B: Lowly deleterious, C: Moderately deleterious, D: Highly deleterious) in continental refugium populations (during the last glacial maxima) versus continental recolonized populations. Recolonized populations were significantly

depleted in minor alleles for all variants categories (Wilcoxon rank sum tests; Neutral: $W = 16067$, $p\text{-value} < 2.2e-16$, effect size = 0.512, considered large; Lowly deleterious: $W = 15998$, $p\text{-value} < 2.2e-16$, effect size = 0.508, considered large; Moderately deleterious: $W = 15997$, $p\text{-value} < 2.2e-16$, effect size = 0.508, considered large; Highly deleterious: $W = 15151$, $p\text{-value} < 2.2e-16$, effect size = 0.453, considered large). However, this is solely due to the higher genetic diversity of refugium populations and this depletion disappeared when we correct for individual genetic diversity (by dividing the count of minor allele by the individual number of polymorphic sites) (figure S9, panels A to D). Concerning the R'_{XY} ratios (refugium/recolonized), it was significantly higher than one for neutral minor alleles, possibly reflecting the lack of neutral genetic diversity in recolonized populations. On the contrary, R'_{XY} ratios were not significantly different from one for all types of deleterious minor alleles (panel E), indicating that recolonized continental populations are as enriched in deleterious minor alleles compared to intergenic minor variants. Panels F to I show the number of homozygous minor alleles per variants category (F: Neutral; G: Lowly deleterious, H: Moderately deleterious; I: Highly deleterious) in refugium populations versus recolonized populations. Recolonized populations were also significantly enriched in homozygous minor alleles for all variants categories (Wilcoxon rank sum tests; Neutral: $W = 12796$, $p\text{-value} = 9.64e-10$, effect size = 0.300, considered small; Lowly deleterious: $W = 12364$, $p\text{-value} = 3.005e-08$, effect size = 0.272, considered small; Moderately deleterious: $W = 12540$, $p\text{-value} = 7.688e-09$, effect size = 0.283, considered small; Highly deleterious: $W = 11056$, $p\text{-value} = 0.0001399$, effect size = 0.187, considered small). Similarly to what we observed for the count of minor alleles, this depletion disappeared when we divided the number of homozygous minor alleles by the individual genetic diversity for all categories of alleles (figure S9, panels E to H). Concerning the R'^2_{XY} ratios (refugium/recolonized), recolonized populations were less enriched than refugium populations in minor homozygous alleles for neutral as well as lowly and mildly deleterious variants but the ratio was not significantly different from one for highly deleterious variants (panel J).

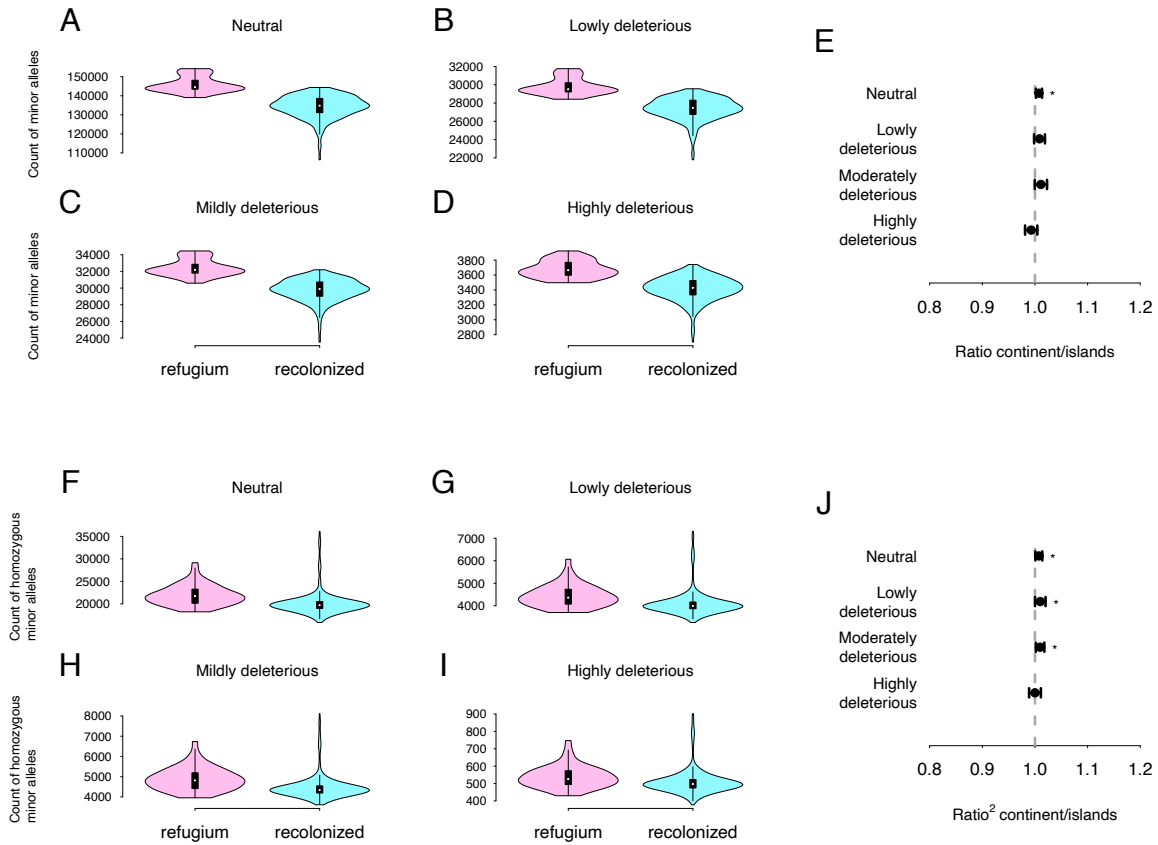


Figure 5: Distribution (among individuals) of minor alleles in continental refugium populations versus recolonized (after the last glacial maximum) populations. Minor allele effects were estimated with *SNPeff*. **A:** Count of neutral minor alleles. **B:** Count of lowly deleterious minor alleles. **C:** Count of moderately deleterious minor alleles. **D:** Count of highly deleterious minor alleles. **E:** R'_{XY} ratio of minor alleles in refugium populations compared to recolonized populations. $R'_{XY} < 1$ indicates that recolonized populations are enriched in minor alleles of the focal category compared to their enrichment for supposedly neutral introns alleles. On the contrary, $R'_{XY} > 1$ indicates that recolonized populations are more depleted in minor alleles of the focal category compared to their enrichment for introns alleles. **F:** Count of homozygous neutral minor alleles. **G:** Count of homozygous lowly deleterious minor alleles. **H:** Count of homozygous moderately deleterious minor alleles. **I:** Count of homozygous highly deleterious minor alleles. **J:** R'^2_{XY} ratio of minor alleles in refugium populations compared to recolonized populations compared to the baseline enrichment for supposedly neutral introns alleles. $R'^2_{XY} < 1$ indicates that recolonized populations are enriched in homozygous minor alleles of the focal category compared to their enrichment in homozygous introns alleles. On the contrary, $R'^2_{XY} > 1$ indicates that recolonized populations are more depleted in homozygous minor alleles of the focal category compared to their enrichment in homozygous introns alleles. For violin plots, refugium populations are shown in pink and recolonized populations in light blue.

Discussion

In this study, we examine the inbreeding status and origin of barn owls across Europe and compare continental and island populations. Our findings indicate that island populations have higher F_{HBD} compared to continental populations and the inbreeding of island individuals is largely a result of ancient coalescence events since 1) F_{HBD} is higher

than F_{AS} which indicates that the populations mostly practice random mating but are still enriched for identical-by-descent (IBD) segments and 2) homozygous-by-descent (HBD) segment distributions demonstrate that island populations exhibit higher numbers of small fragments indicating more ancient coalescence events [63, 61]. These results are consistent with the rank of the estimated N_e , as well as previous studies which showed that the population from CT showed higher inbreeding compared to CY, AE, GR and IS, mostly practiced random mating ($F_{IS} = -0.018$) and was enriched in small ROHs [45] and that both populations from the Canary islands are less heterozygous than PT and MA and mostly practice random mating ($F_{IS} = -0.015$ and -0.030) [12]. Typically, island populations are more inbred due to their smaller size, isolation, and colonization often by a small number of individuals [40]. Interestingly, the GB and IR populations had more HBD segments for the same fraction of genome being autozygous compared to other island populations indicating that even though the fraction of genome within HBD segments is the same, their fragments are on average smaller and closer to the HBD segments distributions from continental populations. In addition, their number of minor alleles count matched the values estimated from continental populations better than values estimated from other islands populations. Their colonization is fairly recent compared to the other islands (it occurred after the last glacial maxima) [44]. In addition, the islands are the largest and exhibit homogeneous structure, consistent with an history of small N_e [44].

The majority of individuals from continental populations were sampled in Switzerland (CH), whose pedigree revealed family structure. This was further confirmed by the plot comparing F_{AS} and F_{HBD} : most CH individuals were below the one-to-one line. Additionally, we identified 14 inbred individuals that were the result of mating events between closely related parents in accordance with the pedigree. The France (FR) population showed a higher mean F_{HBD} and an enrichment in the HBD segment class due to recent coalescence events, and a lower N_e estimate. Inbreeding and N_e estimates were probably biased because one individual (out of five) was strongly inbred, resulting in a high mean F_{HBD} and recent HBD segment sum of lengths. We observed unexpected results for the population from Georgia (GE) which displayed very low N_e and negative F_{AS} . This is most likely due to these samples being full-siblings and therefore sharing more alleles than unrelated individuals, which will influence the estimation of π and N_e . As a result, their F_{AS} is also affected: the sampled population's average allele matching is exceptionally high, but because their parents are not closely related, they do not share

many more homozygous sites than expected under random mating, which causes a low allele matching score multiplied by a very high relatedness average which leads to a negative F_{AS} .

Consequently, with the current samples, we cannot draw any conclusions about what the effective population size estimate as well as the origin of inbreeding in this population. However, the remaining analyses such as F_{HBD} estimation as well as the count of minor alleles should not be biased. Compared to more northern continental populations, Portugal (PT), Morocco (MA) and Israel (IS) had higher N_e estimates and lower HBD segments sum of lengths. This is because these populations are the largest refugia [11]. Despite Italy was also identified as a refugium, its size is smaller, which may explain the smaller N_e and enrichment in small HBD segments. There was little difference in inbreeding levels between recolonized continental populations and refugium continental populations, which is likely due to constant and strong gene flow between both groups, as shown by the low F_{ST} (0.047) among populations [11].

The N_e values we estimated were too high. We believe that this is due to different factors. Firstly, the mutation rate we used was estimated in the collared flycatcher (*Ficedula albicollis*) and with only one family [60]. Our decision was based on the fact that we did not have an estimate of the mutation rate for barn owls. While the absolute N_e values we estimated cannot be used to draw any conclusions, relative comparisons between our populations should still be valid. The high values we obtained are probably also influenced by the strong connectivity among all our populations (the maximum F_{ST} is between CY and PT and is equal to 0.102 [11]). It is therefore likely that neighboring populations are exchanging large numbers of migrants per generation. It has been shown that strong gene flow can inflate populations-specific N_e estimation to the metapopulation N_e value [65]. Furthermore, the AE's very high N_e estimate can also be attributed to the fact that this population barely differs from the large continental GR population ($F_{ST} = 0.014$ [45]). Finally, CY showed a particularly high N_e compared to other islands. This is in agreement with previous studies, who determined that CY is the most diverse island in the Mediterranean Sea and has the strongest gene flow with the mainland [45].

The probability (averaged among individuals) that a genomic region is HBD varied along the genome but was generally quite low. This is especially true when compared to another species where the same probability was estimated: a Soay sheep population [62]. It is anticipated, however, that barn owl populations are less inbred than the Soay sheep which

is a small isolated population with high relatedness [62]. We further showed that, in barn owl, the probability of a region being HBD was strongly correlated with the number of genes in the region: regions with high gene density had a lower probability of being HBD. Therefore, there is a constraint on regions with high gene density to have lower homozygosity probabilities. This is in agreement with inbreeding depression theory, which predicts that homozygosity at a coding site will result in the expression of recessive deleterious alleles, thereby reducing the fitness of individuals [8]. As a result, selection will act to reduce homozygosity in these regions if they contain recessive deleterious alleles. There is no evidence that the high probability of belonging to an HBD segment at the beginning of Super-Scaffold 3 is biologically significant. The high probability of being HBD could be an artifact of the higher coverage of this region, which was twice as much than the rest of the genome (in both our data and the genome assembly). Indeed, higher coverage increases the chance of finding homozygosity, which in turn, increases the probability that a region is HBD.

We showed that most inbreeding is due to small effective population sizes rather than mating between closely related individuals. According to [25], purging solely due to a reduction in population size has been shown to occur only when population size is small enough for drift to increase homozygosity and cause recessive deleterious alleles to be expressed in the majority of individuals [8], without being too small to prevent drift from overcoming selection.

Our findings demonstrate that island populations were enriched in all types of minor alleles in terms of absolute numbers, probably because of drift and suggesting that selection has been less efficient at removing deleterious alleles in island populations. In addition, island populations were similarly enriched for all types of intragenic variants compared to intergenic mutations. This indicates that selection has also been less efficient at removing deleterious alleles in the smaller island populations for all types of deleterious mutations. It is interesting to note that even though it is not significantly different from one, the average ratio of continental and island populations for highly deleterious mutations is slightly higher than one. However, even if it was significant, it would be far from the values which have been reported as evidence for purging in the wild in mountain gorillas ($R'_{XY} = 0.8$) [68] and alpine ibex ($R'_{XY} = 0.525$) [27]. We should, however, note that the bottlenecks of these species were extremely severe compared to what occurred during the colonization of islands by barn owls: the mountain gorilla

population size was estimated at around 800 individuals in 2015 [68] while the Ibex alpine Swiss population has been reintroduced from only 100 individuals with little subsequent gene flow [26]. We hypothesise that the islands populations we are studying are too large to allow for purging. Finally, we note that the same type of enrichment for all categories of deleterious alleles was also observed in humans out-of-Africa expansion [47, 52] but, similarly to what we report in this study, no sign of purging of highly deleterious alleles was detected [16].

Interestingly, all deleterious categories of minor alleles were depleted in recolonized continental populations compared to refugium populations. However, this depletion disappeared when we considered individual genetic diversity which is known to be higher in refugium [11]. Recolonized populations were significantly less depleted in neutral minor alleles compared to intergenic minor alleles (but the ratio was very close to 1) which probably reflects their lower genetic diversity. This pattern was, however, not observed for all types of deleterious mutations, which suggests that selection has been as efficient at removing deleterious alleles for all these categories. Overall, the depletion of minor alleles in recolonized populations goes against what has been reported in the literature [29, 52, 59]. However, this could be explained because the continental populations were recolonized long ago and are not at the extremities of range expansion anymore. In addition, there were multiple source populations for Europe mainland recolonization [11] which could reduce the expansion load [52]. Finally, recolonized continental populations have constant gene flow with the source populations (here the refugium), thus reducing the possibility that rare alleles will rise to high frequencies [51]. Finally, we want to stress that the different $SNPEff$ categories we used could be biased. Indeed, it is very hard to merge different types of mutations into lowly, mildly or highly deleterious variants and be confident about their effect. Some authors have preferred to use more precise categories of variants such as Loss of functions variants (as proxy for highly deleterious mutations), synonymous mutations (as proxy for neutral or lowly deleterious variants) or non-synonymous mutations (as proxy for mildly deleterious variants) [51, 36].

Since we used the global minor allele as a proxy for the derived allele, our analyses of enrichment in deleterious alleles may be biased. Ideally, we should redo the SNP-calling with an outgroup and infer the derived state for as many sites as possible. There will still be many sites for which it is impossible to infer the ancestral and derived alleles (when

both alleles are present in the outgroup) and this would only enable us to analyze a subset of the variants.

Conclusion

In this study, we examined the inbreeding status and the HBD segments landscape of barn owl populations throughout Europe. Compared to continental populations, island populations are more inbred, and the inbreeding primarily results from a small effective population size rather than recent consanguinity. We show that the probability a region is autozygous diminishes with the number of genes present in this region. Finally, we show that in comparison to continental populations, island populations are enriched in all deleterious categories of minor alleles reflecting the lower efficiency of selection at removing deleterious alleles in smaller populations and an absence of purging.

References

- [1] S. Antoniazza, R. Kanitz, S. Neuenschwander, R. Burri, A. Gaigher, A. Roulin, and J. Goudet. Natural selection in a postglacial range expansion: the case of the colour cline in the European barn owl. *Molecular Ecology*, 23(22):5508–5523, 2014.
- [2] A.R. Bertrand, N.K. Kadri, L. Flori, M. Gautier, and T. Druet. *RZooRoH*: An R package to characterize individual genomic autozygosity and identify homozygous-by-descent segments. *Methods in Ecology and Evolution*, 10(6):860–866, 2019.
- [3] A.M. Bolger, M. Lohse, and B. Usadel. *Trimmomatic*: a flexible trimmer for illumina sequence data. *Bioinformatics*, 30(15):2114–2120, 2014.
- [4] R. Burri, S. Antoniazza, A. Gaigher, A-L. Ducrest, C. Simon, The European Barn Owl Network, L. Fumagalli, J. Goudet, and A. Roulin. The genetic basis of color-related local adaptation in a ring-like colonization around the Mediterranean. *Evolution*, 70(1):140–153, 2016.
- [5] F.C. Ceballos, S. Hazelhurst, D.W. Clark, G. Agongo, G. Asiki, P.R. Boua, F.X. Gómez Olivé, F. Mashinya, S. Norris, J.F. Wilson, and M. Ramsay. Autozygosity influences cardiometabolic disease-associated traits in the AWI-Gen sub-Saharan African study. *Nature communications*, 11(1):5754, 2020.
- [6] F.C. Ceballos, P.K. Joshi, D.W. Clark, M. Ramsay, and J.F. Wilson. Runs of homozygosity: windows into population history and trait architecture. *Nature Reviews Genetics*, 19(4):220–234, 2018.
- [7] C.C. Chang, C.C. Chow, L.C. Tellier, S. Vattikuti, S.M. Purcell, and J.J. Lee. Second-generation *PLINK*: rising to the challenge of larger and richer datasets. *GigaScience*, 4(1):7, 2015.
- [8] D. Charlesworth and J.H. Willis. The genetics of inbreeding depression. *Nature Reviews Genetics*, 10(11):783–796, 2009.
- [9] P. Cingolani, A. Platts, L.L. Wang, M. Coon, T. Nguyen, L. Wang, S.J. Land, X. Lu, and D.M. Ruden. A program for annotating and predicting the effects of single nucleotide polymorphisms, *SnpEff*: SNPs in the genome of drosophila melanogaster strain w¹⁸⁸² iso-2; iso-3. *Fly*, 6(2):80–92, 2012.
- [10] J.F. Crow. *Population Theory: Mathematical Topics in Population Genetics*. (ed. Kojima, K. I.). *Springer-Verlag*, pages 128–177, 1970.
- [11] T. Cumer, A.P. Machado, G. Dumont, V. Bontzorlos, R. Ceccherelli, M. Charter, K. Dichmann, N. Kassinis, R. Lourenço, F. Manzia, H-D. Martens, L. Prévost, M. Rakovic, I. Roque, F. Siverio, A.

- Roulin, and J. Goudet. Landscape and Climatic Variations Shaped Secondary Contacts amid Barn Owls of the Western Palearctic. *Molecular Biology and Evolution*, 39(1):msab343, 2022.
- [12] T. Cumer, A.P. Machado, F. Siverio, S.I. Cherkaoui, I. Roque, R. Lourenço, M. Charter, A. Roulin, and J. Goudet. Genomic basis of insularity and ecological divergence in barn owls (*Tyto alba*) of the Canary Islands. *Heredity*, 129(5):281–294, 2022.
- [13] P. Danecek, A. Auton, G. Abecasis, C.A. Albers, E. Banks, M.A. DePristo, R.E. Handsaker, G. Lunter, G.T. Marth, S.T. Sherry, G. McVean, R. Durbin, and 1000 Genomes Project Analysis Group. The variant call format and *VCFTools*. *Bioinformatics*, 27(15):2156–2158, 2011.
- [14] P. Danecek, J.K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M.O. Pollard, A. Whitwham, T. Keane, S.A. McCarthy, R.M. Davies, and H. Li. Twelve years of *SAMtools* and *BCFtools*. *GigaScience*, 10(2):giab008, 2021.
- [15] S.B. Day, E.H. Bryant, and L.M. Meffert. The influence of variable rates of inbreeding on fitness, environmental responsiveness, and evolutionary potential. *Evolution; international journal of organic evolution*, 57(6):1314–1324, 2003.
- [16] R. Do, D. Balick, H. Li, I. Adzhubei, S. Sunyaev, and D. Reich. No evidence that selection has been less effective at removing deleterious mutations in Europeans than in Africans. *Nature Genetics*, 47(2):126–131, 2015.
- [17] T. Druet and M. Gautier. A model-based approach to characterize individual inbreeding at both global and local genomic scales. *Molecular Ecology*, 26(20):5820–5841, 2017.
- [18] T. Druet and M. Gautier. A hidden Markov model to estimate homozygous-by-descent probabilities associated with nested layers of ancestors. *Theoretical Population Biology*, 145:38–51, 2022.
- [19] N. Dussex, H.E. Morales, C. Grossen, L. Dalén, and C. van Oosterhout. Purging and accumulation of genetic load in conservation. *Trends in Ecology & Evolution*, 38(10):961–969, 2023.
- [20] A. Eyre-Walker and P.D. Keightley. The distribution of fitness effects of new mutations. *Nature Reviews Genetics*, 8(8):610–618, 2007.
- [21] D.S. Falconer and M. Falconer. *Introduction to Quantitative Genetics*. Textbook Publishers, 2003.
- [22] G.F. Ficetola, F. Mazel, and W. Thuiller. Global determinants of zoogeographical boundaries. *Nature Ecology & Evolution*, 1(4):0089, 2017
- [23] A. García-Dorado. Understanding and Predicting the Fitness Decline of Shrunk Populations: Inbreeding, Purging, Mutation, and Standard Selection. *Genetics*, 190(4):1461–1476, 2012.
- [24] S.X. Ge, D. Jung, and R. Yao. *ShinyGO*: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics*, 36(8):2628–2629, 2020.
- [25] S. Glémin. How are deleterious mutations purged ? Drift versus non-random mating. *Evolution*, 57(12):2678–2687, 2003.
- [26] C. Grodinsky and M. Stuwe. The reintroduction of the alpine ibex to the swiss alps. *Smithsonian*, 18(9):68, 1987.
- [27] C. Grossen, F. Guillaume, L.F. Keller, and D. Croll. Purging of highly deleterious mutations through severe bottlenecks in Alpine ibex. *Nature Communications*, 11(1):1001, 2020.
- [28] P.W. Hedrick and A. García-Dorado. Understanding Inbreeding Depression, Purging, and Genetic Rescue. *Trends in Ecology & Evolution*, 31(12):940–952, 2016.
- [29] B.M. Henn, L.R. Botigué, S. Peischl, I. Dupanloup, M. Lipatov, B.K. Maples, A.R. Martin, S. Musharoff, H. Cann, M.P. Snyder, L. Excoffier, J.M. Kidd, and C.D. Bustamante. Distance from sub-Saharan Africa predicts mutational load in diverse human genomes. *Proceedings of the National Academy of Sciences*, 113(4), 2016.
- [30] G.M. Hewitt. The genetic legacy of the Quaternary ice ages. *Nature*, 405(6789):907–913, 2000.
- [31] G.M. Hewitt. Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, 68(1-2):87–112, 1999.
- [32] G.M. Hewitt. Mediterranean Peninsulas: The Evolution of Hotspots. In Frank E. Zachos and Jan Christian Habel, editors, *Biodiversity Hotspots: Distribution and Protection of Conservation Priority Areas*, pages 123–147. Springer Berlin Heidelberg, 2011.
- [33] J.R. Homburger, A. Moreno-Estrada, C.R. Gignoux, D. Nelson, E. Sanchez, P. Ortiz-Tello, B.A.

- Pons-Estel, E. Acevedo-Vasquez, P. Miranda, C.D. Langefeld, S. Gravel, M.E. Alarcón-Riquelme, and C.D. Bustamante. Genomic Insights into the Ancestry and Demographic History of South America. *PLOS Genetics*, 11(12):e1005602, 2015.
- [34] C.D. Huber, B.Y. Kim, C.D. Marsden, and K.E. Lohmueller. Determining the factors driving selective effects of new non-synonymous mutations. *Proceedings of the National Academy of Sciences*, 114(17):4465–4470, 2017.
- [35] J. Huisman, L.E.B. Kruuk, P.A. Ellis, T. Clutton-Brock, and J.M. Pemberton. Inbreeding depression across the lifespan in a wild mammal population. *Proceedings of the National Academy of Sciences*, 113(13):3585–3590, March 2016.
- [36] E. Humble, M.A. Stoffel, K. Dicks, A.D. Ball, R.M. Gooley, J. Chuvén, R. Pusey, M. Al Remeithi, K-P. Koepfli, B. Pukazhenthi, H. Senn, and R. Ogden. Conservation management strategy impacts inbreeding and mutation load in scimitar-horned oryx. *Proceedings of the National Academy of Sciences*, 120(18):e2210756120, 2023.
- [37] M. Kardos, E.E. Armstrong, S.W. Fitzpatrick, S. Hauser, P.W. Hedrick, J.M. Miller, D.A. Tallmon, and W.C. Funk. The crucial role of genome-wide genetic variation in conservation. *Proceedings of the National Academy of Sciences*, 118(48):e2104642118, 2021.
- [38] M. Kardos, Y. Zhang, K.M. Parsons, Y.A. H.Kang, X. Xu, X. Liu, C.O. Matkin, P. Zhang, E.J. Ward, M.B. Hanson, C. Emmons, M.J. Ford, G. Fan, and S. Li. Inbreeding depression explains killer whale population dynamics. *Nature Ecology & Evolution*, 7(5):675–686, 2023.
- [39] R. Kassen and T. Bataillon. Distribution of fitness effects among beneficial mutations before selection in experimental populations of bacteria. *Nature Genetics*, 38(4):484–488, 2006.
- [40] L.F. Keller, and D.G. Waller. Inbreeding effects in wild populations. *Trends in Ecology & Evolution*, 17(5):230–241, 2002.
- [41] T.T. Kibota and M. Lynch. Estimate of the genomic mutation rate deleterious to overall fitness in *E. coll.* *Nature*, 381(6584):694–696, 1996.
- [42] M. Kirkpatrick and P. Jarne. The effects of a bottleneck on inbreeding depression and the genetic load. *The American Naturalist*, 155(2):154–167, 2000.
- [43] H. Li and R. Durbin. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*, 25(14):1754–1760, 2009.
- [44] A.P. Machado, T. Cumer, C. Iseli, E. Beaudoin, A-L. Ducrest, M. Dupasquier, N. Guex, K. Dichmann, R. Lourenço, J. Lusby, H-D. Martens, L. Prévost, D. Ramsden, A. Roulin, and J. Goudet. Unexpected post-glacial colonisation route explains the white colour of barn owls (*tyto alba*) from the british isles. *Molecular Ecology*, 31(2):482–497, 2021.
- [45] A.P. Machado, A. Topaloudis, T. Cumer, E. Lavanchy, V. Bontzorlos, R. Ceccherelli, M. Charter, N. Kassinis, P. Lymberakis, F. Manzia, A-L. Ducrest, M. Dupasquier, N. Guex, A. Roulin, and J. Goudet. Genomic consequences of colonisation, migration and genetic drift in barn owl insular populations of the eastern Mediterranean. *Molecular Ecology*, 31(5):1375–1388, 2022.
- [46] K. Martikainen, A. Sironen, and P. Uimari. Estimation of intrachromosomal inbreeding depression on female fertility using runs of homozygosity in Finnish Ayrshire cattle. *Journal of Dairy Science*, 101(12):11097–11107, 2018.
- [47] R.C. McCoy and J.M. Akey. Patterns of deleterious variation between human populations reveal an unbalanced load. *Proceedings of the National Academy of Sciences*, 113(4):809–811, 2016.
- [48] R. McQuillan, A-L. Leutenegger, R. Abdel-Rahman, C.S. Franklin, M. Pericic, L. Barac-Lauc, N. Smolej-Narancic, B. Janicijevic, O. Polasek, A. Tenesa, A.K. MacLeod, S.M. Farrington, P. Rudan, C. Hayward, V. Vitart, I. Rudan, S.H. Wild, M.G. Dunlop, A.F. Wright, H. Campbell, and J.F. Wilson. Runs of Homozygosity in European Populations. *The American Journal of Human Genetics*, 83(3):359–372, 2008.
- [49] N. Nakatsuka, P. Moorjani, N. Rai, B. Sarkar, A. Tandon, N. Patterson, G.S. Bhavani, K.M. Girisha, M.S. Mustak, S. Srinivasan, A. Kaushik, S.A. Vahab, S.M. Jagadeesh, K. Satyamoorthy, L. Singh, D. Reich, and K. Thangaraj. The promise of discovering population-specific disease-associated genes in South Asia. *Nature Genetics*, 49(9):1403–1407, 2017.
- [50] V. Narasimhan, P. Danecek, A. Scally, Y. Xue, C. Tyler-Smith, and R. Durbin. *BCFtools/RoH*: a hidden

- Markov model approach for detecting autozygosity from next-generation sequencing data. *Bioinformatics*, 32(11):1749–1751, 2016.
- [51] SF. Nigenda-Morales, M Lin, PG. Nuñez-Valencia, C.C. Kyriazis, A.C. Beichman, J.A. Robinson, A.P. Ragsdale, J. Urbán, F.I. Archer, L. Viloría-Gómora, M.J. Pérez-Álvarez, E. Poulin, K.E. Lohmueller, A. Moreno-Estrada, and R.K. Wayne. The genomic footprint of whaling and isolation in fin whale populations. *Nature Communications*, 14(1):5465, 2023.
- [52] S. Peischl, I. Dupanloup, M. Kirkpatrick, and L. Excoffier. On the accumulation of deleterious mutations during range expansions. *Molecular Ecology*, 22(24):5972–5982, 2013.
- [53] J.K. Pritchard. Are Rare Variants Responsible for Susceptibility to Complex Diseases? *The American Journal of Human Genetics*, 69(1):124–137, 2001.
- [54] J.E. Pryce, M. Haile-Mariam, M.E. Goddard, and B.J. Hayes. Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle. *Genetics Selection Evolution*, 46(1):71, 2014.
- [55] S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M.A.R. Ferreira, D. Bender, J. Maller, P. Sklar, P.I.W. de Bakker, M.J. Daly, and P.C. Sham. *PLINK*: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, 81(3):559–575, 2007.
- [56] P. Ralph and G. Coop. The Geography of Recent Genetic Ancestry across Europe. *PLoS Biology*, 11(5):e1001555, 2013.
- [57] P. Rastas. *Lep-MAP3*: robust linkage mapping even for low-coverage whole genome sequencing data. *Bioinformatics*, 33(23):3726–3732, 2017.
- [58] J.A. Robinson, C. Brown, B.Y. Kim, K.E. Lohmueller, and R.K. Wayne. Purging of Strongly Deleterious Mutations Explains Long-Term Persistence and Absence of Inbreeding Depression in Island Foxes. *Current Biology*, 28(21):3487–3494.e4, 2018.
- [59] Q. Rougemont, T. Leroy, E.B. Rondeau, B. Koop, and L. Bernatchez. Allele surfing causes maladaptation in a Pacific salmon of conservation concern. *PLOS Genetics*, 19(9):e1010918, 2023.
- [60] L. Smeds, A. Qvarnström, and H. Ellegren. Direct estimate of the rate of germline mutation in a bird. *Genome Research*, 26(9):1211–1218, 2016.
- [61] D. Speed and D.J. Balding. Relatedness in the post-genomic era: is it still useful? *Nature Reviews Genetics*, 16(1):33–44, 2015.
- [62] M.A. Stoffel, S.E. Johnston, J.G. Pilkington, and J.M. Pemberton. Genetic architecture and lifetime dynamics of inbreeding depression in a wild mammal. *Nature Communications*, 12(1):2972, 2021.
- [63] E.A. Thompson. Identity by Descent: Variation in Meiosis, Across Genomes, and in Populations. *Genetics*, 194(2):301–326, 2013.
- [64] J. Wang, W.G. Hill, D. Charlesworth, and B. Charlesworth. Dynamics of inbreeding depression due to deleterious mutations in small populations: mutation parameters and inbreeding rate. *Genetical Research*, 74(2):165–178, 1999.
- [65] R.S. Waples and P.R. England. Estimating Contemporary Effective Population Size on the Basis of Linkage Disequilibrium in the Face of Migration. *Genetics*, 189(2):633–644, 2011.
- [66] B.S. Weir and J. Goudet. A unified characterization of population structure and relatedness. *Genetics*, 206(4):2085–2103, 2017.
- [67] Y. Willi, M. Fracassetti, S. Zoller, and J. van Buskirk. Accumulation of Mutational Load at the Edges of a Species Range. *Molecular Biology and Evolution*, 35(4):781–791, 2018.
- [68] Y. Xue, J. Prado-Martinez, P.H. Sudmant, V. Narasimhan, Q. Ayub, M. Szpak, P. Frandsen, Y. Chen, B. Yngvadottir, D.N. Cooper, M. de Manuel, J. Hernandez-Rodriguez, I. Lobon, H.R. Siegismund, L. Pagani, M.A. Quail, C. Hvilsom, A. Mudakikwa, E.E. Eichler, M.R. Cranfield, T. Marques-Bonet, C. Tyler-Smith, and A. Scally. Mountain gorilla genomes reveal the impact of long-term population decline and inbreeding. *Science*, 348(6231):242–245, 2015.
- [69] Q.S. Zhang, J. Goudet, and B.S. Weir. Rank-invariant estimation of inbreeding coefficients. *Heredity*, 128(1):1–10, 2022.

Chapter IV: Inbreeding depression in the Swiss barn owl (*Tyto alba*) population

Authors: Eléonore Lavanchy^{1,2}, Anna Hewett^{1,2}, Tristan Cumer^{1,2}, Alexandros Topaloudis^{1,2}, Anne-Lyse Ducrest¹, Céline Simon¹, Alexandre Roulin¹ and Jérôme Goudet^{1,2}

¹ Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland

² Swiss Institute of Bioinformatics, University of Lausanne, Lausanne, Switzerland

Chapter status: not submitted

Authors contribution: EL and JG conceptualized the study. A-LD and CS performed the DNA extraction and libraries preparations. EL, TC and AT generated genotypes data. EL performed analyses related to Homozygous-by-descent segments and inbreeding depression analyses in adults. AH performed inbreeding depression analyses in juveniles. EL wrote the manuscript with inputs and feedbacks from all co-authors.

Abstract

Inbreeding and its negative effect on individuals is ubiquitous but examples of its effects in the wild are scarce. Our study combines high coverage and imputed low coverage data on 3,085 barn owls (*Tyto alba*) and quantifies the severity of inbreeding depression in juveniles and adults separately. Inbreeding appears to have a mild effect on juvenile bill length: a fully inbred individual's bill length decreases by 10% compared with a non-inbred individual. There was no association between bill length, nor any other morphological trait, and inbreeding status in adults, suggesting that inbreeding depression varies with age.

Introduction

Inbreeding depression is defined as a decrease in fitness in inbred individuals. The phenomenon has been extensively studied in theory [11] and observed empirically in several domestic species and in humans. For example, increased individual inbreeding coefficients leads to reduced milk production [5, 62, 22, 21] as well as reproductive ability [5, 22, 27, 62] in cattle. Similarly, in humans, inbreeding has been linked to a number of diseases, including Alzheimer's disease [68, 52, 26] and Schizophrenia [40] as well as height reduction [49] and reduced fertility [60, 15]. Hence, given its ubiquity, the quantification of the extent of inbreeding depression in populations as well as understanding its underlying mechanisms has been a priority for the conservation of both wild and captive species.

Two main hypotheses have been proposed to explain the mechanisms underlying inbreeding depression [11]. First, it can be due to the accumulation of (slightly) deleterious recessive alleles in a homozygous state [11]. Second, it can be due to loci with heterozygous advantage for which heterozygous individuals have a greater fitness than homozygous individuals. An important consequence of inbreeding is the increase in autozygosity (i.e. identical-by-descent (IBD) homozygosity) [72, 17, 48], which in both cases leads to a fitness disadvantage. Several (partially) recessive deleterious mutations are found to segregate at low to intermediate frequencies in natural populations [61]. Due to their recessive nature, their effect is only fully expressed when they are homozygous. Therefore, the number of expressed deleterious recessive mutations is positively correlated with the degree of inbreeding of individuals, and it is the accumulation of such mutations that contributes to the detrimental effects of inbreeding. In large populations,

selection maintains deleterious mutations at low frequencies. The effect of drift, however, is greatly increased in small populations and can result in an increase in the frequency of these mutations [39]. There has been significant evidence for the importance of heterozygous advantage in reducing inbred individual fitness in drosophila [11] and rice [46], but it may not be as important in other species, particularly since loci with heterozygous advantage are relatively scarce, and certainly rarer than segregating deleterious recessive mutations [11, 30].

How to quantify the direct effect of inbreeding on individuals' fitness or phenotypes has been widely debated in the literature [41, 54, 74, 36, 1, 8]. A common approach consists in regressing the individual inbreeding coefficient on the trait using a (generalized) linear model and including covariates that might affect the trait (e.g. sex [66, 32, 15, 37] age [66, 15, 37] or relatedness structure such as in the second chapter of this thesis and in [55, 66, 56, 49]). Other approaches such as maximum likelihood estimation have also been proposed [54] but are rarely used in practice.

To evaluate the strength of inbreeding depression, it is necessary to first estimate the level of individual inbreeding, which is quantified by inbreeding coefficients (denoted by F). Following the initial work of Sewall Wright with pedigree data [72] and with the advances of sequencing technologies, various methods have been proposed to assess the inbreeding status of individuals: i) comparing the expected and observed heterozygosities [10, 63], ii) studying the correlation between uniting gametes [73, 76] iii) quantifying the expected allele-sharing of individuals [28, 69] or iv) identifying the fraction of genome within IBD segments [50, 38, 23].

Until recently, there was no consensus on which is the best approach to quantify the effects of inbreeding on a given trait [1, 8, 74, 35, 36], with different studies presented conflicting results. Yengo *et al.* found that F_{UNI} (a coefficient based on the correlation of uniting gametes) was the most accurate for quantifying inbreeding depression with human data [74]. On the contrary, several studies found that F_{ROH} (an inbreeding coefficient quantifying the fraction of genome within IBD segments) provided more accurate estimates of the strength of inbreeding depression using simulated data [41, 35, 54]. Based on simulated data including various populations' sizes and histories, Caballero *et al.* [8] concluded that population size determines the most appropriate inbreeding coefficient for measuring inbreeding depression strength. This conclusion was further confirmed using empirical cattle data [1]. Indeed, F_{UNI} correlates better with

homozygosity at rare alleles while F_{ROH} correlates better with homozygosity at intermediate alleles. Selection is highly efficient in large populations (such as humans) and deleterious alleles are maintained at low frequencies. Consequently, F_{UNI} provides a more accurate estimation of homozygosity for deleterious alleles in large populations. However, as the population size decreases, the influence of drift becomes stronger and can equalize the influence of selection. This results in deleterious alleles reaching intermediate frequencies and, in this case, F_{ROH} correlates better with deleterious allele homozygosity.

As mentioned above, several studies have been conducted on inbreeding depression in populations that have been thoroughly studied with large data sets, such as domestic animals and humans [60, 41, 67, 5]. Detecting inbreeding depression in the wild, however, involves additional challenges: unless the strength of inbreeding depression is exceptionally high, a large number of samples would be required to detect inbreeding depression accurately (as shown in chapter II of this thesis and in [40, 8]) which is hardly feasible in natural populations, especially if they are endangered. In addition, large amount of phenotypic data (correctly linked to the genotypic data) are required which is also complicated to obtain in natural populations. Recent advancements in monitoring, sampling and sequencing resources have made this possible and recent studies showed an association between inbreeding and different traits such as breeding success decrease in deer [32], lower survival in both wolves [45] and whales [37], population growth decrease in ibex [7] and survival and growth vigor decrease in plants [75, 3, 33].

The barn owl (*Tyto alba*) is a nocturnal bird of prey widely distributed throughout Europe. Since the last glacial maximum, barn owls have colonized Switzerland from refuges in Portugal and Greece [18]. For over 30 years, a subset the Swiss population has been subjected to a long-term study, which involves collecting pedigree data, blood samples, and morphological data. A precise description of the data collection process was described in [25]. As birds can fly in and out of the study area quite easily, the pedigree of the birds is quite shallow and our current estimations suggest that 60% of the population is composed of migrants (i.e. birds born outside the study area). Overall, this population and the large amount of data collected through the years are ideal for studying inbreeding effects on phenotypes in the wild.

In this paper, we use a combination of high and low coverages whole genome re-sequencing data of 3,085 Swiss barn owls to study their inbreeding status as well as the

effect of inbreeding depression on different morphological traits such as tarsus and bill length and mass in both adults and juveniles separately. We found significant inbreeding depression for some traits on young individuals but not on adults.

Material and Methods

Sampling, sequencing and genotyping

We sequenced a total of 3,085 barn owls (*Tyto alba*) sampled as part of a long-term study in South-Western Switzerland. 358 of these individuals were sequenced at high coverage (14X on average), herein referred to as the ‘reference panel’, and used to phase and impute the remaining 2,768 individuals sequenced at low coverage (on average 1.91X). Annex I and II provide detailed descriptions of high-coverage genotyping and phasing respectively, while Annex III provides precise descriptions of low-coverage genotypes generation. Briefly, DNA was extracted from blood and tissue samples with the DNeasy Blood & Tissue kit (Qiagen, Hilden, Germany) for both high and low coverages samples. Whole-genome re-sequencing was performed at the Lausanne Genomic Technologies Facility (GTF, University of Lausanne, Switzerland) using Illumina HiSeq 2500 PE high-throughput sequencing. We then trimmed and aligned raw reads to the reference genome [47] with `Trimomatic` (v0.39) [6] and `BWA-MEM` (v.0.7.17) [44].

Following read alignment, the pipeline differs for high and low coverage individuals. For the reference panel, we used the ‘Base quality score recalibration’ (BQSR) method as suggested in GATK best practices for non-model species. We then called haplotypes using GATK’s (v4.2.6.1) *HaplotypeCaller* method and performed diploid joint-calling [59]. We then filtered variants following GATK ‘best practices’ (QD < 2.0, FS > 60.0, MQ < 40.0, MQRankSum < -12.5, ReadPosRankSum < -8.0, SOR > 3.0) and performed an additional ‘mappability’ filter (described in [16] and adapted from [these guidelines](#)) to mask regions for which the mapping quality (to the reference genome) was low. We then filtered on individuals’ depth using `BCFTools` (v.1.15.1) [19]: we set any variant with individual depth < 5 and > the individual mean depth plus three times its standard deviation to missing and filtered out any site with minor allele count (MAC) < 5. We also filtered out any site with more than 5% (of individuals) missing data and any non-bi-allelic variant. We used such hard filters because these high coverage genotypes are then used for

phasing and imputation of the low-coverage genotypes. Consequently, we need them to be of high quality.

We then used these high-coverage genotypes to phase and impute the low-coverage genotypes. The procedure is described in detail in Annex III but to summarize, we first obtained genotypes likelihood for the low coverage individuals with `BCFTOOLS` (v.1.15.1) `mpileup` and `call` [43] at the varying position in the high coverage reference panel only. We then used the software `GLIMPSE` (v.1) [65] to phase and impute these variant positions. We used the first version of the software rather than the last available version because it is recommended for smaller reference panels on the website. In brief, in `GLIMPSE`, low-coverage individuals' haplotypes are reconstructed as a combination of the closest haplotypes in the high-coverage reference panel and of the reconstructed low-coverage haplotypes. We show in Annex III that this method results in reliable genotype calls (correlation of genotypes for 32 individuals sequenced both at low and high coverages was 0.98 for 50,000 randomly selected SNPs).

HBD segments identification

Homozygous-by-descent (HBD) segments were called as described in the previous chapter. Briefly, we used genetic positions (i.e. we took recombination rate into account) and used the `RZOOROH` package (v.0.3.1) [23, 4, 24] with 13 HBD classes and 1 non-HBD class with rates (R) of 2, 4, 8, 16, 32, 64, 128, 256, 502, 1024, 2048, 4096 and 8192 for the HBD classes and 8192 for the non-HBD class [23]. For HBD segments distributions, we considered the 'most probable HBD segments' estimated by `RZOOROH` Viterbi algorithm and estimated the mean (among individuals) sum of HBD segments lengths falling into the different HBD classes. We considered a HBD class as identical-by-descent (IBD) if the rate was lower or equal to 1024 (i.e. if the coalescence event occurred up to 512 generations ago).

Estimating inbreeding coefficients

We averaged the probability of belonging to a HBD segment in order to estimate the HBD segments-based inbreeding coefficient: F_{HBD} (with the `cumhbd` function from the `RZOOROH` package with a T value of 1024). This value means that only HBD segments coalescing less than 512 generations ago are considered as IBD. We also estimated F_{AS} ,

an inbreeding coefficient relying on the average allele-sharing between individuals (scaled by the mean allele-sharing between individuals) [76] estimated as follow:

$$F_{ASj} = \frac{\sum_{l=1}^L (A_{jl} - A_{Sl})}{\sum_{l=1}^L (1 - A_{Sl})} \quad (1)$$

where A_{jl} corresponds the state of the two alleles an individual j carries at locus l : one for homozygous and 0 for heterozygous and A_{Sl} is the average allele sharing proportion at locus l for pairs of individuals $j, k, j \neq k$.

For inbreeding depression modelling, we used another inbreeding coefficient: F_{UNI} initially described in [73] and modified in [76] which was shown to be the best inbreeding coefficients for estimating inbreeding depression in large populations (in chapter II of this thesis and in [1, 8, 74]). F_{UNI} measures the correlation between uniting gametes and the modified version we used (here and in chapter II) gives less weight to rare alleles:

$$F_{UNIj}^w = \frac{\sum_{l=1}^L (x_{jl}^2 - (1+2p_l)x_l + 2p_l^2)}{\sum_{l=1}^L 2p_l(1-p_l)} \quad (2)$$

with $x_{jl} \in \{0, 1, 2\}$ being the minor allele count of individual j at locus l and p_l being the derived allele frequency at the same locus.

Estimating heritability and inbreeding depression

We used the R package `MCMCglmm` for all statistical modelling. For each trait we modelled adults and juveniles separately under the assumption that parameters affecting the traits will differ between life stages. Firstly, we built standard generalised linear mixed models with fixed and random effects tailored to each trait and life stage (see table 1 and supplementary material equations S1 – S6 for specific fixed and random effects). Next, to determine the effect of individuals' inbreeding coefficient on the trait (i.e inbreeding depression) we fitted F_{UNI} as a fixed effect (equations in supplementary material). Finally, to estimate heritability of traits we used the 'animal model' framework which uses a pedigree to relate individuals to their additive genetic values. We calculated narrow sense heritability as the proportion of phenotypic variance explained by additive genetic variance, as such:

$$h^2 = V_A / V_P \quad (3)$$

Where V_A and V_P are the additive genetic variance and phenotypic variance respectively. All animal models were run with the standard following structure:

$$y = X\beta + Z_1a + Z_r u_r + e \quad (4)$$

Where, y is a vector of the trait measurements, X is an incidence matrix relating individual measures to the vector of fixed effects β . Z_1 and Z_r are incidence matrices relating individual measures to additive genetic and remaining random effects respectively; a is the relatedness matrix estimated from the pedigree; u_r is a vector of additional random effects; and finally, e is a vector of residual effects. For juveniles, the animal model was incorporated into the existing model of inbreeding depression described above. In contrast, for adults we estimated inbreeding depression and heritability in separate models. This is because F_{UNI} was not significant for adults and greatly reduced the sample size (our pedigree contained many more adults than we sequenced) which was not the case for juveniles. Table 1, below, describes the parameters included as fixed and random effects for each phenotypic trait and life stage. Details are also given on the number of MCMC iterations and the burn in used which we optimised through inspection of model convergence in trace plots and the effective sample size.

Table 1: Description of the parameters used in the `MCMCglmm` models used for heritability estimation. F_{UNI} is the inbreeding coefficient. Sex is the sex of the individual. Rank is the rank of birth of the individual in the clutch. For juveniles, the age of an individual (in days) was included in the model with a Gompertz growth function as $y = Asym * e^{(-b_2 * b_3^x)}$, where x is the age in days and the 3 remaining parameters were estimated for each trait using the `SSgompertz` function in R. Pedigree represents the relatedness matrix estimated from the pedigree (i.e. the ARM). Individual ID represents a unique ID per individual to account for repeated measures across individuals. Observer ID is the person who took the measurement. Clutch ID is the ID of the clutch in which the bird was born and raised. Birth year is the year in which the individual was born. For adults, LastChickHatched refers to a binary parameter indicating whether the last chick of the clutch already hatched or not (i.e if the females is still sitting on the eggs). All models described in this table are fully described in equations S1 - S6.

Trait	Life Stage	Fixed Effects	Random Effects	$n_{iterations}$	n_{burnin}
Bill length	Juveniles	F_{UNI} Sex Rank Age	Pedigree Individual ID Clutch ID Observer ID Birth Year	250,000	50,000
Bill length	Adults	Sex	Pedigree Individual ID Clutch ID Observer ID Birth Year	1,000,000	50,000
Mass	Juveniles	F_{UNI} Sex Rank Age	Pedigree Individual ID Clutch ID Observer ID Birth Year	250,000	50,000
Mass	Adults	Sex LastChickHatched Sex:LastChickHatched	Pedigree Individual ID Clutch ID Observer ID Birth Year	1,000,000	50,000
Tarsus length	Juveniles	F_{UNI} Sex Rank Age	Pedigree Individual ID Clutch ID Observer ID Birth Year	250,000	50,000
Tarsus length	Adults	Sex Rank	Pedigree Individual ID Clutch ID Observer ID Birth Year	1,000,000	50,000

Results

Figure 1 shows the F_{HBD} distribution of the population (panel A). The majority of the population estimates of F_{HBD} are < 0.1 , indicating less than 1% of strongly inbred individuals. Panel B presents F_{HBD} versus F_{AS} estimates; each point corresponds to one individual. F_{AS} can be seen as a proxy of how much the alleles within an individual are associated at random. If $F_{AS} = 0$ the individual comes from parents which kinship equals the average kinship of the population (i.e. the parents are probably unrelated so the individual is probably not inbred). On the contrary, if $F_{AS} > 0$, the parents are more related than the average relatedness in the population (i.e. the parents are most likely related and the individual is probably inbred). Finally, if an individual shows a $F_{AS} < 0$, it means that their parents are less related than the average of the population (i.e. the parents are likely unrelated and the individual could even be admixed). Consequently, comparing F_{AS} and F_{HBD} can indicate where inbreeding comes from. Panel B shows that the few individuals with $F_{HBD} > 0.1$ also have F_{AS} higher than zero, suggesting that their inbreeding is mostly due to mating between related parents and not to population's small effective size. In addition, most individuals are below the identity line (i.e. $F_{AS} > F_{HBD}$) suggesting we have structure in our data. Indeed, when $F_{HBD} < F_{AS}$, it indicates that the individual is not autozygous at many sites but that its parents are more similar than the average of the population which indicates that there is population structure in our data set (the individual is part of a subgroup of the population). This is not surprising as our data set includes many families. Nevertheless, among the individuals below the line, several have high F_{AS} and F_{HBD} , indicating that we sampled few families which are inbred. Finally, we have some individuals with highly negative F_{AS} values (and $F_{HBD} < 0.05$). In comparison to the general population, these individuals' parents are less related. Panel C shows the probability that a variant is autozygous along the genome. For most regions of the genome, this probability is low. Nevertheless, we have a very high peak at the beginning of Super-Scaffold 3 on around 400Kb. In this region, we did not identify any particular gene or gene enrichment. Other smaller peaks are often found at the beginning or end of super-scaffolds.

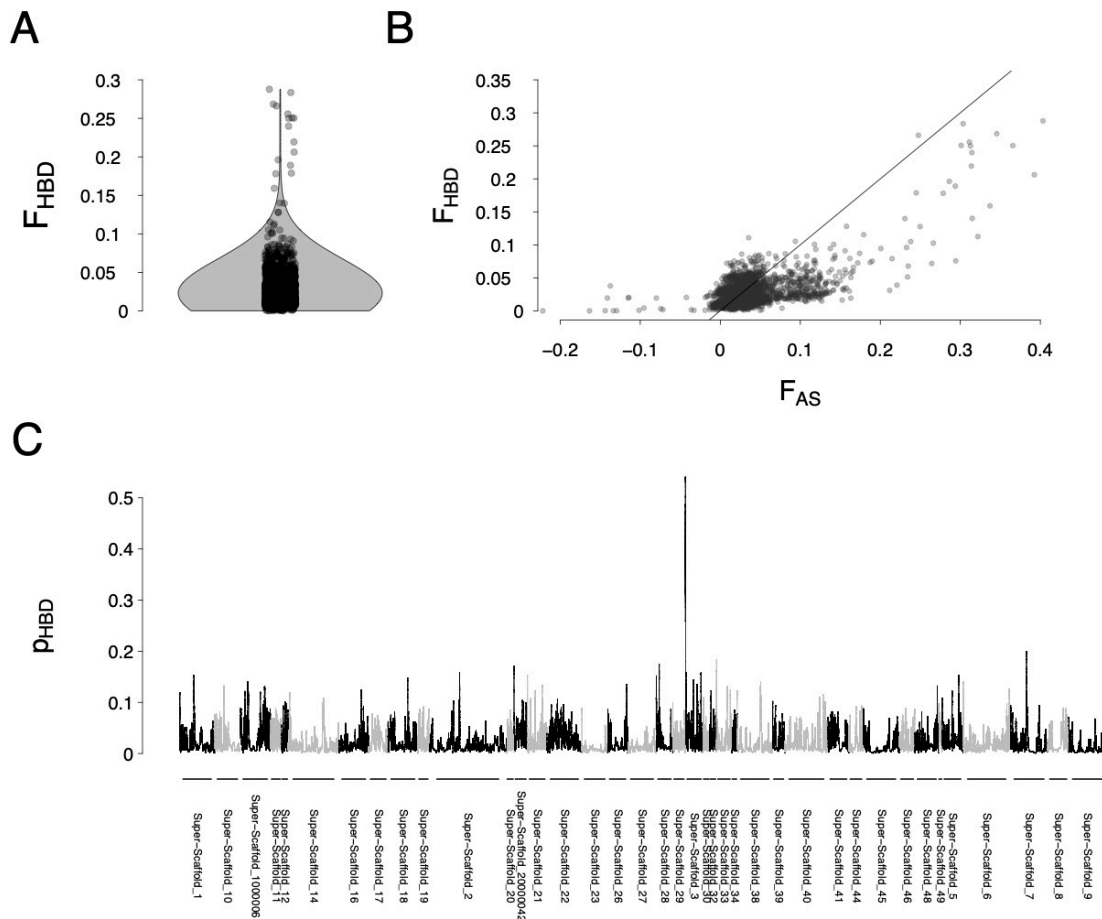


Figure 1: Inbreeding quantification. **A:** F_{HBD} distribution **B:** F_{HBD} according to F_{AS} **C:** Probability that a marker belongs to an HBD segment (p_{HBD}) along the genome, averaged in 100Kb sliding windows with 20Kb overlap. Alternating colors correspond to a change in super-scaffold.

Figure 2 illustrates different representations of the distribution of HBD segments among the 3,085 samples. Panel A shows the total number of HBD segments for each individual according to the total length of these segments. First, most individuals have a small number of HBD segments (less than 300) and second, most individuals follow a linear trajectory. However, few individuals (16) have fewer segments than expected according to their sum of all HBD segments. These individuals are the result of matings between parents who are closely related and carry long HBD segments. In contrast, inbred individuals with many segments are more likely to derive from matings between relatively ancient relatives. Panel B shows the distribution of segments falling into the different HBD classes. This demonstrates that most individuals carry HBD segments that date from ancient coalescent events (256 and 512 generations on average) and very few HBD segments that date from recent coalescence events (less than eight generations ago). Finally, panel C shows the HBD segments carried by the three individuals with the most

and least inbred genomes. The least inbred individuals carry between one and six HBD segments all smaller than 1Mb. In contrast, the most inbred individual (M047568) has one HBD segment exceeding 26Mb. It should be noted, however, that most of the HBD segments in M047568 are much smaller (median length of HBD segments in M047568: 312Kb; mean length of HBD segments in M047568: 1Mb).

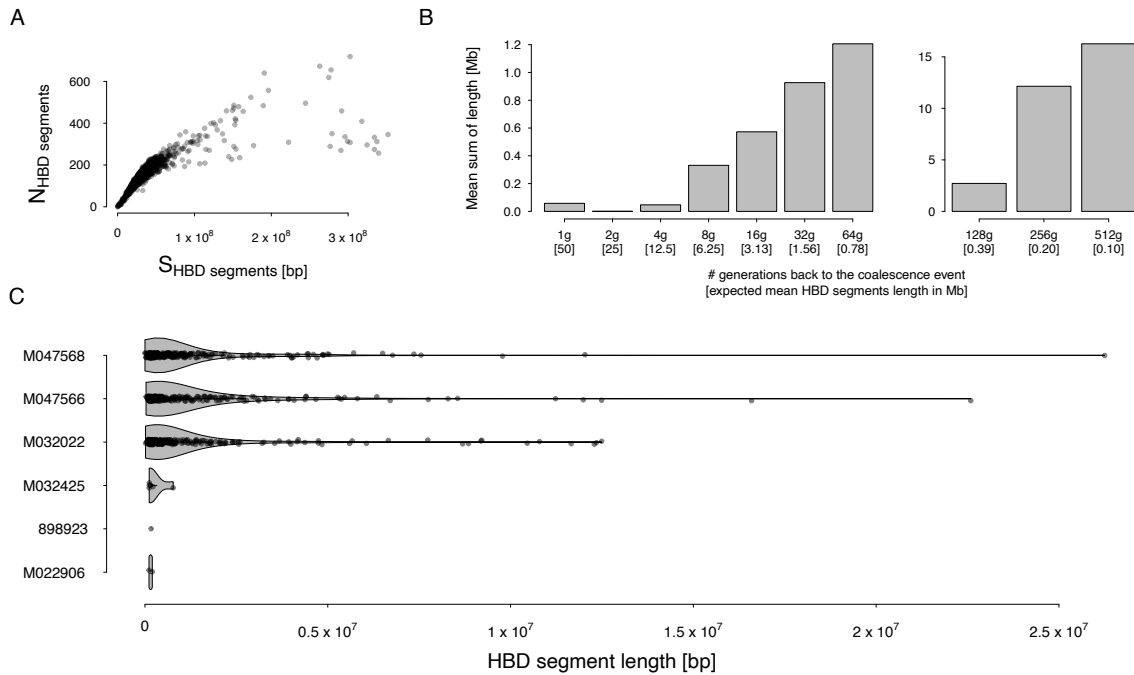


Figure 2: HBD segments distribution. **A:** Number of HBD segments according to sum of HBD segments per individual **B:** HBD segments distribution **C:** HBD segments distributions for the three less and more inbred individuals in the data set.

Figure 3 shows the different morphological traits variance partitioning among the random effects of the MCMCg1mm models for both juveniles and adults. Heritability values extracted from the same models are presented in table 2. Since it had no significant effect and was not available for most adults, F_{UNI} was not included in the animal models for adults. The mean (across MCMC iterations) heritability (V_A/V_P) of adults was always larger than that of nestlings. Similarly, larger heritabilities were observed from the distribution of the V_A estimates from the different iterations of the MCMC chain (figure S1). The variation between iterations was quite large especially for tarsus length in adults (table 2 and figure S1). Concerning other random effects, the clutch in which the birds was born and grew ($V_{\text{Clutch ID}}$), which represents all the environment the bird was exposed to as a juvenile, such as its parents and how good hunters they were, was important for all three traits in juveniles, but only for tarsus length in adults. Similarly, the bird's birth

year ($V_{Birth\ Year}$) was always more influential in juveniles than in adults. With the exception of mass measurement in juveniles, who took the measurement ($V_{Observer\ ID}$) had only a reduced effect. Next, we observe almost no effect of individual measurement ($V_{Individual\ ID}$ the variance among repeated measures of the same individual) suggesting that the repeatability of our measurements is high (repeatability measured as the intraclass correlation coefficient (ICC) in adults is 0.667 (CI = [0.639; 0.694]) for bill length, 0.594 (CI = [0.554; 0.633]) for mass and 0.801 (CI = [0.768; 0.829]) for tarsus length). Finally, the residuals ($V_{Residuals}$ i.e. the fraction of variance not explained by our model) are the highest for the mass model in adults and for the tarsus length model in nestlings.

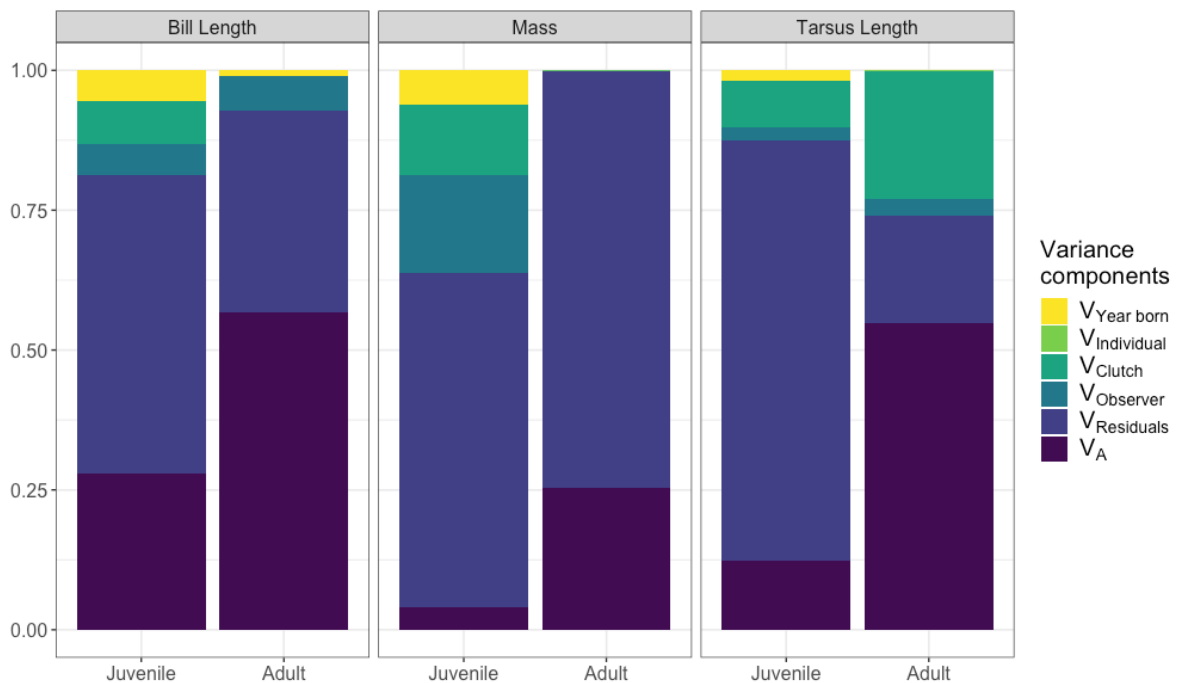


Figure 3: Variance partitioning among the different random effects of our animal models for both nestlings and adults. The inbreeding coefficient F_{UNI} was included in the models for nestlings but not for adults. Birth Year is the year in which the individual was born. Individual ID represents a unique ID per individual to account for repeated measures across individuals. Clutch ID is the ID of the clutch in which the bird was born and raised. Observer ID is the person who took the measurement. Residuals represent the variance not explained by our model. V_A is the variance explained by the pedigree, i.e. the additive genetic variance.

Table 2: Heritability (h^2) estimates for the different traits in both adults and juveniles. Heritability estimates were extracted from MCMCglmm models as additive variance divided by the total phenotypic variance: V_A/V_P . The inbreeding coefficient F_{UNI} was included for juveniles only.

Trait	Life Stage	Mean h^2	95% prediction interval
Bill length	Juveniles	0.26	[0.19; 0.33]
Bill length	Adults	0.50	[0.41; 0.58]
Mass	Juveniles	0.04	[0.00; 0.07]
Mass	Adults	0.21	[0.09; 0.32]
Tarsus length	Juveniles	0.12	[0.07; 0.17]
Tarsus length	Adults	0.50	[0.31; 0.68]

Figure 4 illustrates the inbreeding depression slope estimates (β) for the different phenotypes. Panel A shows morphological traits for both juveniles and adults. Inbreeding depression was detected only for bill length in juveniles (mean $\beta = -17.13$; 95% CI = [-25.7205; -8.3872]). Hence, an individual with an inbreeding coefficient F_{UNI} of 0 has, on average, a bill 1.71 mm longer than an individual with a F_{UNI} of 1. This represents a 10.6% decrease in bill length. Only the β estimate of bill length was negative for adults, but all CI overlapped with zero. Adults had particularly large CIs for both mass and tarsus length. It should be noted that although the CI around β estimates of mass and tarsus length overlapped with zero in juveniles, they were negative. Panel B shows fitness-related traits for adults only. We did not detect inbreeding depression on any of these traits (the β estimates are positive and all CI overlapped with 0).

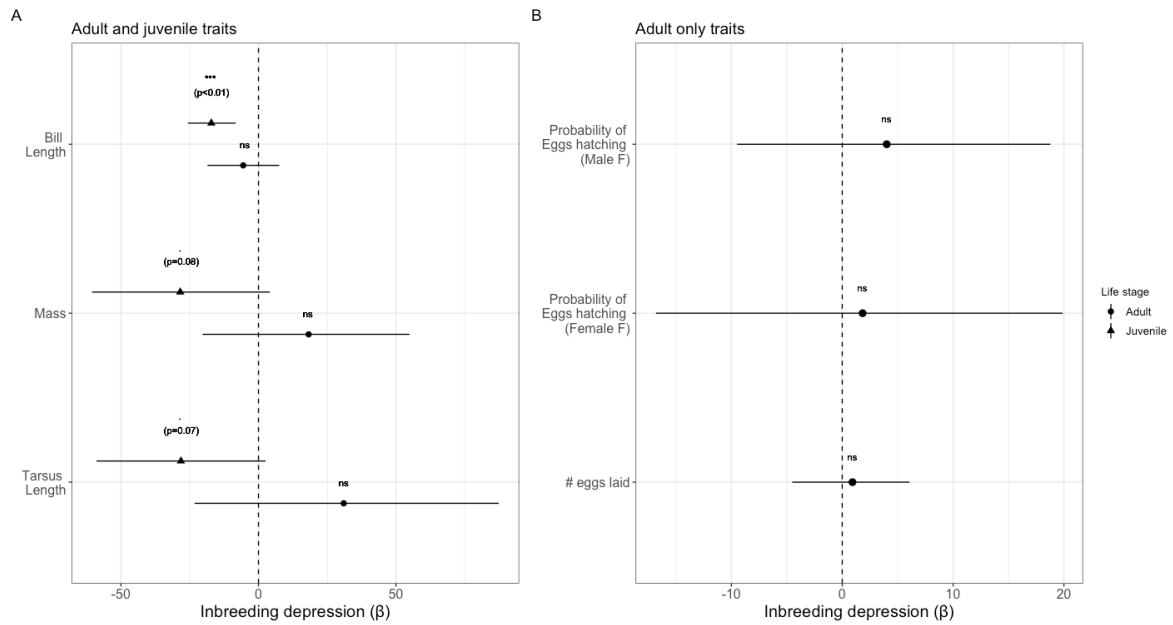


Figure 4: Inbreeding depression strength (β) estimates estimated with F_{UNI} . **A:** for both juveniles and adults in the three morphological traits. Models correspond to equations S16, S18, S20, S22, S24 and S26 in supplementary material. **B:** Yearly fitness-related traits for adults only. Number of eggs laid is for females only while probability that an egg hatches is for both parents. Models correspond to equations S28 and S32 in supplementary material. For both panels, sample sizes represent the number of observations.

We ran models with and without pedigree information in both juveniles and adults and compared inbreeding depression quantification obtained with F_{UNI} and F_{HBD} (figure S2). We found that i) there is hardly any difference between the models including or excluding the pedigree and that ii) the CI for β estimated with F_{HBD} are always larger than the estimates from F_{UNI} , especially in adults.

Furthermore, we would like to emphasize that we included rank (in birth) in the inbreeding depression model for tarsus length in adults because it had a significant influence on tarsus length. The majority of adults, however, are born outside of our study system, and their rank is unknown. Therefore, it greatly reduced the sample size and the range of F_{UNI} covered. This resulted in β shifting from zero (when rank is not taken into account) to positive (when rank is included in the model) but did not change the significance. Comparison between both models (with and without rank can be found in figure S3).

Discussion

In this paper, we first characterize the status and origin of inbreeding in the Swiss barn owl population and look for evidence of inbreeding depression. We first show that the majority of our samples have low levels of close inbreeding due to both their low fraction of genome within HBD segments [50, 9] low HBD segment total numbers [9, 53] and average length [42, 9]. In addition, we demonstrate that the few highly inbred individuals are clustered into families and are inbred because their parents are closely related. In contrast, we identified 16 outbred individuals with negative F_{AS} values (< -0.05). The parents or grandparents of all but one of these individuals are unknown, and they may be either migrants from neighboring populations or offspring of a migrant and a local. We found that the probability of a marker being autozygous varied along the genome with an inflated value at the very beginning of Super Scaffold 3. It is likely that this peak is due to an assembly problem. In both the data used to generate the assembly and our data, we observe an inflated coverage in this particular area (the coverage in this region is on average twice as high as the coverage in the rest of the genome). This probably indicates an amplification of this particular region. As for the rest of the genome, the variation we observe may be the result of drift, selection, or recombination [58, 31]. However, it is more likely to be drift or selection, since we have considered the local recombination rate in our model.

We found that heritability was always higher in adults than in juveniles. There is also evidence that heritability increases with age in other species, such as sheep for weight and hindleg length [70, 71]. In swans, heritability follows a quadratic curve: low in nestlings, almost zero in young adults, and high in older individuals [12]. It could be due to senescence and the accumulation of mutations acting at late life stages [51, 12]. However, increased heritability in adults is not always due to an increased additive genetic variance (V_A) but could also be due to a decrease in phenotypic variance (V_P), i.e. due to increased parental effects in juveniles [70]. Indeed, juveniles are more dependent on their environment because their parents feed them, whereas adults are solely responsible for feeding themselves and surviving. We found that the clutch in which the bird was born and raised, and hence the parental investment, had no effect on adults' bill length and mass. On the contrary, it always explained at least 10% of the variance in juveniles. Adult tarsus length had 23% of its variance explained by clutch ID, which suggests that the juvenile environment is more important for tarsus growth than bill

growth. This conclusion is further supported by the fact that the rank at birth of an individual significantly impacts the length of the tarsus in adults. At the juvenile stage, a barn owl's rank determines how much food is available. Barn owls lay their eggs asynchronously, and they hatch approximately every two to three days within a clutch. Therefore, the first born is larger and stronger than his siblings and will be able to access the food brought back by his parents more easily [64]. As a result, older siblings are able to feed and grow more rapidly. Bringing enough food for large clutches is only possible if the clutch is small or if the parents are extremely successful hunters. We also demonstrate that juveniles were more affected by the year in which they were born than adults were. The condition of the chicks in the nest is partly dependent on the severity of winters [2] or the summer weather as rainy weather conditions hinders the hunting success of parents [13]. Interestingly, the year of birth had very little effect on tarsus length at both life stages. It could be because the variance of year has already been absorbed by our clutch ID variable. Or it could be an ascertainment bias because all the individuals with small tarsus died regardless of the year they were born and we did not sample them. There was little variance explained by who measured the traits except for mass in juveniles. This is because half of the people used a scale for mass measurement in juveniles while the other half used a pesola. Regarding the fraction of variance not explained by our models, the residuals in adults were the highest for mass. This suggests that we have a poor model for mass. It could be explained partly because we cannot know whether an owl is fed when we weight it, and we estimated that an adult owl gains approximately eight percent of its weight after eating a prey. We also showed that variation among heritability estimates were higher for tarsus length in adults. This is likely due to the smaller sample size for tarsus length ($n = 1,100$ observations for 592 individuals) compared to both bill length and mass ($n = 3,321$ for 1,860 individuals and $n = 1,932$ for 598 individuals respectively). This is because we included the rank (of birth) of individuals in the tarsus length model because its effect was significant (which was not the case for both bill length and mass). Since most of our adults are born outside our study system, we don't know what their rank is and we had to remove many tarsus observations.

We found that inbreeding significantly affects bill length in juveniles but not in adults. There is a possibility that this is the result of a smaller sample size and smaller range of F for adults or (and most likely) that these adults actually passed the first selection process: survival after leaving the nest, resulting in a lesser effect of inbreeding on

survivors due to the fact that individuals who were less fit did not survive [57, 66, 14]. A similar negative effect of inbreeding on bill length has previously been reported in house sparrows [56]. Concerning the other morphological traits, inbreeding has a negative effect on both mass and tarsus length in juveniles, but the effect is not significant. It may be that our models are less accurate for both of these traits (the residuals are large). In addition, the confidence intervals are just overlapping with zero and increasing the sample size could lead to their reduction and a significant effect of inbreeding depression. However, we cannot make any conclusions and certainly not that it would be significant if our models were more accurate or if we had more observations. Among adults, both mass and tarsus length estimates are positive, but non-significant and harbor large confidence intervals. It is probably due to the large residuals of the model for mass. For tarsus length, it could be a result of the smaller sample size and the smaller range of F_{UNI} covered. β estimates for fitness-related traits in adults are close to zero, with very large CI. For the fitness-related traits, it is unexpected as we expect life-history traits such as probability that an egg hatches to be more sensitive to inbreeding depression [56, 20]. However, it is likely because our sample sizes are very small (we only used 56 clutches for the probability that an egg hatches). However, just like before we cannot conclude that we would detect inbreeding depression with more observations.

The results of our study add to the growing body of literature where recent studies have documented more and more instances of inbreeding depression in the wild [66, 32, 37, 29, 56]. For instance, a 10% increase in F_{ROH} has been shown to decrease survival by 60% in Soay sheep [66]. Similarly, Kardos and collaborators [37] found that an increase of 14% in F_{ROH} decreased the probability of killer whales' survival to 40 years by 71%. Other studies showed that lifetime breeding success was reduced by 84% for individuals from half-sibling mating in red deer [32] and by 88% for a 9% increase in homozygosity in the helmet honeyeater [29]. Our study reports relatively low inbreeding effects compared to what was found in these studies. It is probably because what we are reporting is a morphological trait, which usually has little bearing on fitness compared with survival or breeding success. This lesser effect of inbreeding on morphological traits compared to survival or breeding success has been reported previously in house sparrow [56] and mentioned in previous studies [39, 20]. A possible explanation is that morphological traits may have a small individual effect on fitness but add up over an individual's lifetime to influence reproductive success [56, 34].

Finally, in the supplementary material, we compare both F_{UNI} and F_{HBD} for inbreeding depression estimation. Even though we don't know the true value of β here, the small CI around β , coupled with the fact that the owls we are studying come from a population with large effective size and numbers, suggests that F_{UNI} is likely more reliable than F_{HBD} in our population [1, 8].

Conclusion

The Swiss barn owl population consists primarily of individuals with low inbreeding coefficients, however, there are a few individuals that are the result of mating between closely related parents. We show that increased levels of inbreeding lead to negative effects such as reduced bill length in juveniles, but not in adults, suggesting that inbreeding depression varies depending on the individual's life stage.

References

- [1] S.W. Alemu, N.K. Kadri, C. Harland, P. Faux, C. Charlier, A. Caballero, and T. Druet. An evaluation of inbreeding measures using a whole-genome sequenced cattle pedigree. *Heredity*, 126(3):410–423, 2021.
- [2] R. Altwegg, A. Roulin, M. Kestenholz, and L. Jenni. Demographic effects of extreme winter weather in the barn owl. *Oecologia*, 149(1):44–51, August 200
- [3] F. Angeloni, N.J. Ouborg, and R. Leimu. Meta-analysis on the association of population size and life history with inbreeding depression in plants. *Biological Conservation*, 144(1):35–43, 2011.
- [4] A.R. Bertrand, N. K. Kadri, L. Flori, M. Gautier, and T. Druet. *RZooRoH*: An R package to characterize individual genomic autozygosity and identify homozygous-by-descent segments. *Methods in Ecology and Evolution*, 10(6):860–866, 2019.
- [5] D.W. Bjelland, K.A. Weigel, N. Vukasinovic, and J.D. Nkrumah. Evaluation of inbreeding depression in Holstein cattle using whole-genome SNP markers and alternative measures of genomic inbreeding. *Journal of Dairy Science*, 96(7):4697–4706, 2013.
- [6] A.M. Bolger, M. Lohse, and B. Usadel. Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics*, 30(15):2114–2120, 2014.
- [7] C. Bozzuto, I. Biebach, S. Muff, A.R. Ives, and L.F. Keller. Inbreeding reduces long-term growth of Alpine ibex populations. *Nature Ecology & Evolution*, 3(9):1359–1364, 2019.
- [8] A. Caballero, B. Villanueva, and T. Druet. On the estimation of inbreeding depression using different measures of inbreeding from molecular markers. *Evolutionary Applications*, 14(2):416–428, 2021.
- [9] F.C. Ceballos, P.K. Joshi, D.W. Clark, M. Ramsay, and J.F. Wilson. Runs of homozygosity: windows into population history and trait architecture. *Nature Reviews Genetics*, 19(4):220–234, 2018.
- [10] C.C. Chang, C.C. Chow, L.C. Tellier, S. Vattikuti, S.M. Purcell, and J.J. Lee. Second-generation *PLINK*: rising to the challenge of larger and richer datasets. *GigaScience*, 4(1):7, 2015.
- [11] D. Charlesworth and J.H. Willis. The genetics of inbreeding depression. *Nature Reviews Genetics*, 10(11):783–796, 2009.
- [12] A. Charmantier, C. Perrins, R.H. McCleery, and B.C. Sheldon. Age-dependent genetic variance in a life-history trait in the mute swan. *Proceedings of the Royal Society B: Biological Sciences*, 273(1583):225–232, 2006.

- [13] A. Chausson, I. Henry, B. Almasi, and A. Roulin. Barn Owl (*Tyto alba*) breeding biology in relation to breeding season climate. *Journal of Ornithology*, 155(1):273–281, 2014.
- [14] N. Chen, E.J. Cosgrove, R. Bowman, J.W. Fitzpatrick, and A.G. Clark. Genomic Consequences of Population Decline in the Endangered Florida Scrub-Jay. *Current Biology*, 26(21):2974–2979, 2016.
- [15] D.W. Clark, Y. Okada, K.H.S. Moore, and *et. al* Associations of autozygosity with a broad range of human phenotypes. *Nature Communications*, 10(1):4957, 2019.
- [16] H. Corval, T. Cumer, A. Topaloudis, A. Roulin, and J. Goudet. Where and when local adaptation happens: lessons from the european barn owl (*tyto alba*). *bioRxiv*, 2023.
- [17] J.F. Crow and M. Kimura. *An Introduction to Population Genetics Theory*. Blackburn Press, 2009.
- [18] T. Cumer, A.P. Machado, G. Dumont, V. Bontzorlos, R. Ceccherelli, M. Charter, K. Dichmann, N. Kassinis, R. Lourenço, F. Manzia, H-D. Martens, L. Prévost, M. Rakovic, I. Roque, F. Siverio, A. Roulin, and J. Goudet. Landscape and Climatic Variations Shaped Secondary Contacts amid Barn Owls of the Western Palearctic. *Molecular Biology and Evolution*, 39(1):msab343, 2022.
- [19] P. Danecek, J.K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M.O. Pollard, A. Whitwham, T. Keane, S.A. McCarthy, R.M. Davies, and H. Li. Twelve years of *SAMtools* and *BCFtools*. *GigaScience*, 10(2):giab008, 2021.
- [20] M.A. DeRose and D.A. Roff. A comparison of inbreeding depression in life-history and morphological traits in animals. *Evolution*, 53(4):1288–1292, 1999
- [21] H.P. Doekes, P. Bijma, R.F. Veerkamp, G. De Jong, Y.C.J. Wientjes, and J.J. Windig. Inbreeding depression across the genome of Dutch Holstein Friesian dairy cattle. *Genetics Selection Evolution*, 52(1):64, 2020.
- [22] H.P. Doekes, R.F. Veerkamp, P. Bijma, G. De Jong, S.J. Hiemstra, and J.J. Windig. Inbreeding depression due to recent and ancient inbreeding in Dutch Holstein–Friesian dairy cattle. *Genetics Selection Evolution*, 51(1):54, 2019.
- [23] T. Druet and M. Gautier. A model-based approach to characterize individual inbreeding at both global and local genomic scales. *Molecular Ecology*, 26(20):5820–5841, 2017.
- [24] T. Druet and M. Gautier. A hidden Markov model to estimate homozygous-by-descent probabilities associated with nested layers of ancestors. *Theoretical Population Biology*, 145:38–51, 2022.
- [25] C. Frey, C. Sonnay, A. Dreiss, and A. Roulin. Habitat, breeding performance, diet and individual age in swiss barn owls (*tyto alba*). *Journal of Ornithology*, 152:279–290, 2011.
- [26] M. Ghani, C. Reitz, R. Cheng, B.N. Vardarajan, G. Jun, C. Sato, A. Naj, R. Rajbhandary, L-S. Wang, O. Valladares, C-F. Lin, E.B. Larson, N.R. Graff-Radford, D. Evans, P.L. De Jager, P.K. Crane, J.D. Buxbaum, J.R. Murrell, T. Raj, N. Ertekin-Taner, M. Logue, C.T. Baldwin, R.C. Green, L.L. Barnes, L.B. Cantwell, M.D. Fallin, R.C.P. Go, P.A. Griffith, T.O. Obisesan, J.J. Manly, K.L. Lunetta, M.I. Kamboh, O.L. Lopez, D.A. Bennett, H. Hendrie, K.S. Hall, A.M. Goate, G. S. Byrd, W.A. Kukull, T.M. Foroud, J.L. Haines, L.A. Farrer, M.A. Pericak-Vance, J.H. Lee, G.D. Schellenberg, P. St. George-Hyslop, R. Mayeux, and E. Rogaeva. Association of Long Runs of Homozygosity With Alzheimer Disease Among African American Individuals. *JAMA Neurology*, 72(11):1313, 2015.
- [27] O. González-Recio, E. López De Maturana, and J.P. Gutiérrez. Inbreeding Depression on Female Fertility and Calving Ease in Spanish Dairy Cattle. *Journal of Dairy Science*, 90(12):5744–5752, 2007.
- [28] J. Goudet, T. Kay, and B.S. Weir. How to estimate kinship. *Molecular Ecology*, 27(20):4121–4135, 2018.
- [29] K.A. Harrisson, M.J.L. Magrath, J.D.L. Yen, A. Pavlova, N. Murray, B. Quin, P. Menkhorst, K.A. Miller, K. Cartwright, and P. Sunnucks. Lifetime Fitness Costs of Inbreeding and Being Inbred in a Critically Endangered Bird. *Current Biology*, 29(16):2711–2717.e4, 2019.
- [30] P.W. Hedrick. What is the evidence for heterozygote advantage selection? *Trends in Ecology & Evolution*, 27(12):698–704, 2012.
- [31] A.M. Hewett, M.A. Stoffel, L. Peters, S.E. Johnston, and J.M. Pemberton. Selection, recombination and population history effects on runs of homozygosity (ROH) in wild red deer (*Cervus elaphus*). *Heredity*, 130(4):242–250, 2023.
- [32] J. Huisman, L.E.B. Kruuk, P.A. Ellis, T. Clutton-Brock, and J.M. Pemberton. Inbreeding depression

- across the lifespan in a wild mammal population. *Proceedings of the National Academy of Sciences*, 113(13):3585–3590, 2016.
- [33] B.C. Husband and D.W. Schemske. Evolution of the Magnitude and Timing of Inbreeding Depression in Plants. *Evolution*, 50(1):54, 1996.
- [34] H. Jensen, B-E. Saether, T.H. Ringsby, J. Tufto, S.C. Griffith, and H. Ellegren. Lifetime reproductive success in relation to morphology in the house sparrow *Passer domesticus*. *Journal of Animal Ecology*, 73(4):599–611, 2004.
- [35] M. Kardos, G. Luikart, and F.W. Allendorf. Measuring individual inbreeding in the age of genomics: marker-based measures are better than pedigrees. *Heredity*, 115(1):63–72, 2015.
- [36] M. Kardos, P. Nietlisbach, and P.W. Hedrick. How should we compare different genomic estimates of the strength of inbreeding depression? *Proceedings of the National Academy of Sciences*, 115(11), 2018.
- [37] M. Kardos, Y. Zhang, K.M. Parsons, Y.A. H. Kang, X. Xu, X. Liu, C.O. Matkin, P. Zhang, E.J. Ward, M.B. Hanson, C. Emmons, M.J. Ford, G. Fan, and S. Li. Inbreeding depression explains killer whale population dynamics. *Nature Ecology & Evolution*, 7(5):675–686, 2023.
- [38] M. Kardos, M. Åkesson, T. Fountain, Ø. Flagstad, O. Liberg, P. Olason, H. Sand, P. Wabakken, C. Wikenros, and H. Ellegren. Genomic consequences of intensive inbreeding in an isolated wolf population. *Nature Ecology & Evolution*, 2(1):124–131, 2017.
- [39] L.F. Keller, and D.G. Waller. Inbreeding effects in wild populations. *Trends in Ecology & Evolution*, 17(5):230–241, 2002
- [40] M.C. Keller, M.A. Simonson, S. Ripke, B.M. Neale, P. V. Gejman, D.P. Howrigan, S.H. Lee, T. Lencz, D.F. Levinson, P.F. Sullivan, and The Schizophrenia Psychiatric Genome-Wide Association Study Consortium. Runs of Homozygosity Implicate Autozygosity as a Schizophrenia Risk Factor. *PLoS Genetics*, 8(4):e1002656, 2012.
- [41] M.C. Keller, P.M. Visscher, and M.E. Goddard. Quantification of Inbreeding Due to Distant Ancestors and Its Detection Using Dense Single Nucleotide Polymorphism Data. *Genetics*, 189(1):237–249, 2011.
- [42] M. Kirin, R. McQuillan, C.S. Franklin, H. Campbell, P.M. McKeigue, and J.F. Wilson. Genomic Runs of Homozygosity Record Population History and Consanguinity. *PLoS ONE*, 5(11):e13996, 2010.
- [43] H. Li. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, 27(21):2987–2993, 2011.
- [44] H. Li and R. Durbin. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*, 25(14):1754–1760, 2009.
- [45] O. Liberg, H. Andrén, H-C. Pedersen, H. Sand, D. Sejberg, P. Wabakken, M. Åkesson, and S. Bensch. Severe inbreeding depression in a wild wolf *Canis lupus* population. *Biology Letters*, 1(1):17–20, 2005.
- [46] L.J. Luo, Z-K. Li, H.W. Mei, Q.Y. Shu, R. Tabien, D.B. Zhong, C.S. Ying, J.W. Stansel, G.S. Khush, and A.H. Paterson. Overdominant Epistatic Loci Are the Primary Genetic Basis of Inbreeding Depression and Heterosis in Rice. II. Grain Yield Components. *Genetics*, 158(4):1755–1771, 2001.
- [47] A.P. Machado, T. Cumer, C. Iseli, E. Beaudoin, A-L. Ducrest, M. Dupasquier, N. Guex, K. Dichmann, R. Lourenço, J. Lusby, H-D. Martens, L. Prévost, D. Ramsden, A. Roulin, and J. Goudet. Unexpected post-glacial colonisation route explains the white colour of barn owls (*tyto alba*) from the british isles. *Molecular Ecology*, 31(2):482–497, 2021.
- [48] G. Malécot. *Les mathématiques de l'hérédité*. Masson, 1948.
- [49] R. McQuillan, N. Eklund, N. Pirastu, M. Kuningas, B.P. McEvoy, T. Esko, T. Corre, G. Davies, M. Kaakinen, L-P. Lyytikäinen, K. Kristiansson, A.S. Havulinna, M. Gögele, V. Vitart, A. Tenesa, Y. Aulchenko, C. Hayward, A. Johansson, M. Boban, S. Ulivi, A. Robino, V. Boraska, W. Igl, S.H. Wild, L. Zgaga, N. Amin, E. Theodoratou, O. Polasek, G. Girotto, L.M. Lopez Cinzia Sala, J. Lahti, T. Laatikainen, I. Prokopenko, M. Kals, J. Viikari, J. Yang, A. Pouta, K. Estrada, A. Hofman, N. Freimer, N.G. Martin, M. Kähönen, L. Milani, M. Heliövaara, E. Vartiainen, K. Räikkönen, C. Masciullo, J.M. Starr, A.A. Hicks, L. Esposito, I. Kolcíc, S.M. Farrington, B. Oostra, T. Zemunik, H.

- Campbell, M. Kirin, M. Pehlic, F. Faletra, D. Porteous, G. Pistis, E. Widén, V. Salomaa, S. Koskinen, K. Fischer, T. Lehtimäki, A. Heath, M.I. McCarthy, F. Rivadeneira, G.W. Montgomery, H. Tiemeier, A-L. Hartikainen, P.A.F. Madden, P. d'Adamo, N.D. Hastie, U. Gyllensten, A.F. Wright, C.M. Van Duijn, M. Dunlop, I. Rudan, P. Gasparini, P.P. Pramstaller, I.J. Deary, D. Toniolo, J.G. Eriksson, A. Jula, O.T. Raitakari, A. Metspalu, M. Perola, M-R. Jarvelin, A. Uitterlinden, P.M. Visscher, J.F. Wilson on behalf of the ROHgen Consortium. Evidence of Inbreeding Depression on Human Height. *PLoS Genetics*, 8(7):e1002655, 2012.
- [50] R. McQuillan, A-L. Leutenegger, R. Abdel-Rahman, C.S. Franklin, M. Pericic, L. Barac-Lauc, N. Smolej-Narancic, B. Janicijevic, O. Polasek, A. Tenesa, A.K. MacLeod, S.M. Farrington, P. Rudan, C. Hayward, V. Vitart, I. Rudan, S.H. Wild, M.G. Dunlop, A.F. Wright, H. Campbell, and J.F. Wilson. Runs of Homozygosity in European Populations. *The American Journal of Human Genetics*, 83(3):359–372, 2008.
- [51] P.B. Medawar. An unsolved problem of biology. H.K. Lewis and Company, 1952.
- [52] M.A. Nalls, R.J. Guerreiro, J. Simon-Sanchez, J.T. Bras, B.J. Traynor, J.R. Gibbs, L. Launer, J. Hardy, and A.B. Singleton. Extended tracts of homozygosity identify novel candidate genes associated with late-onset Alzheimer's disease. *neurogenetics*, 10(3):183–190, 2009.
- [53] T.N. Nguyen, N. Chen, E.J. Cosgrove, R. Bowman, J.W. Fitzpatrick, and A.G. Clark. Dynamics of reduced genetic diversity in increasingly fragmented populations of florida scrub jays, *aphelocoma coerulescens*. *Evolutionary Applications*, 15(6):1018–1027, 2022.
- [54] P. Nietlisbach, S. Muff, J.M. Reid, M.C. Whitlock, and L.F. Keller. Nonequivalent lethal equivalents: Models and inbreeding metrics for unbiased estimation of inbreeding load. *Evolutionary Applications*, 12(2):266–279, 2019.
- [55] M. Nishio, K. Inoue, S. Ogawa, K. Ichinoseki, A. Arakawa, Y. Fukuzawa, T. Okamura, E. Kobayashi, M. Taniguchi, M. Oe, and K. Ishii. Comparing pedigree and genomic inbreeding coefficients, and inbreeding depression of reproductive traits in Japanese Black cattle. *BMC Genomics*, 24(1):376, 2023.
- [56] A.K. Niskanen, A.M. Billing, H. Holand, I.J. Hagen, Y.G. Araya-Ajoy, A. Husby, B. Rønning, A.M. Myhre, P.S. Ranke, T. Kvalnes, H. Pärn, T.H. Ringsby, S. Lien, B-E. Sæther, S. Muff, and H. Jensen. Consistent scaling of inbreeding depression in space and time in a house sparrow metapopulation. *Proceedings of the National Academy of Sciences*, 117(25):14584–14592, 2020.
- [57] M. Olsson, A. Gullberg, and H. Tegelström. Malformed offspring, sibling matings, and selection against inbreeding in the sand lizard (*Lacerta agilis*). *Journal of Evolutionary Biology*, 9(2):229–242, 1996.
- [58] T.J. Pemberton, D. Absher, M.W. Feldman, R.M. Myers, N.A. Rosenberg, and J.Z. Li. Genomic Patterns of Homozygosity in Worldwide Human Populations. *The American Journal of Human Genetics*, 91(2):275–292, 2012.
- [59] R. Poplin, V. Ruano-Rubio, M.A. DePristo, T.J. Fennell, M.O. Carneiro, G.A. Van der Auwera, D.E. Kling, L.D. Gauthier, A. Levy-Moonshine, D. Roazen, K. Shakir, J. Thibault, S. Chandran, C. Whelan, M. Lek, S. Gabriel, M.J. Daly, B. Neale, D.G. MacArthur, and E. Banks. Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv*, 2018.
- [60] E. Postma, L. Martini, and P. Martini. Inbred women in a small and isolated Swiss village have fewer children. *Journal of Evolutionary Biology*, 23(7):1468–1474, 2010.
- [61] J.K. Pritchard. Are Rare Variants Responsible for Susceptibility to Complex Diseases? *The American Journal of Human Genetics*, 69(1):124–137, 2001.
- [62] J.E. Pryce, M. Haile-Mariam, M.E. Goddard, and B.J. Hayes. Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle. *Genetics Selection Evolution*, 46(1):71, 2014.
- [63] S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M.A.R. Ferreira, D. Bender, J. Maller, P. Sklar, P.I.W. De Bakker, M.J. Daly, and P.C. Sham. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, 81(3):559–575, 2007.
- [64] A. Roulin. Effects of hatching asynchrony on sibling negotiation, begging, jostling for position and within-brood food allocation in the barn owl, *tyto alba*. *Evolutionary Ecology Research*, 6:1083–1098,

2004.

- [65] S. Rubinacci, D.M. Ribeiro, R.J. Hofmeister, and O. Delaneau. Efficient phasing and imputation of low-coverage sequencing data using large reference panels. *Nature Genetics*, 53(1):120–126, 2021.
- [66] M.A. Stoffel, S.E. Johnston, J.G. Pilkington, and J.M. Pemberton. Genetic architecture and lifetime dynamics of inbreeding depression in a wild mammal. *Nature Communications*, 12(1):2972, 2021.
- [67] N.A. Swinford, S.P. Prall, S. Gopalan, C.M. Williams, J. Sheehama, B.A. Scelza, and B.M. Henn. Increased homozygosity due to endogamy results in fitness consequences in a human population. *Proceedings of the National Academy of Sciences*, 120(43):e2309552120, 2023.
- [68] B.N. Vardarajan, D.J. Schaid, C. Reitz, R. Lantigua, M. Medrano, I.Z. Jiménez-Vélazquez, J.H. Lee, M. Ghani, E. Rogaeva, P. St George-Hyslop, and R.P. Mayeux. Inbreeding among Caribbean Hispanics from the Dominican Republic and its effects on risk of Alzheimer disease. *Genetics in Medicine*, 17(8):639–643, 2015.
- [69] B.S. Weir and J. Goudet. A Unified Characterization of Population Structure and Relatedness. *Genetics*, 206(4):2085–2103, 2017.
- [70] A.J. Wilson, J.M. Pemberton, J.G. Pilkington, T.H. Clutton-Brock, D.W. Coltman, and L.E.B. Kruuk. Quantitative genetics of growth and cryptic evolution of body size in an island population. *Evolutionary Ecology*, 21(3):337, 2007.
- [71] A.J. Wilson and D. Réale. Ontogeny of additive and maternal genetic effects: Lessons from domestic mammals. *The American Naturalist*, 167(1):E23–E38, 2006.
- [72] S. Wright. Coefficients of Inbreeding and Relationship. *The American Naturalist*, 56(645):330–338, 1922.
- [73] J. Yang, S.H. Lee, M.E. Goddard, and P.M. Visscher. *GCTA*: A Tool for Genome-wide Complex Trait Analysis. *The American Journal of Human Genetics*, 88(1):76–82, 2011.
- [74] L. Yengo, Z. Zhu, N.R. Wray, B.S. Weir, J. Yang, M.R. Robinson, and P.M. Visscher. Detection and quantification of inbreeding depression for complex traits from SNP data. *Proceedings of the National Academy of Sciences*, 114(32):8602–8607, 2017.
- [75] C. Zhang, P. Wang, D. Tang, Z. Yang, F. Lu, J. Qi, N.R. Tawari, Y. Shang, C. Li, and S. Huang. The genetic basis of inbreeding depression in potato. *Nature Genetics*, 51(3):374–378, 2019.
- [76] Q.S. Zhang, J. Goudet and B.S. Weir. Rank-invariant estimation of inbreeding coefficients. *Heredity*, 128(1):1–10, 2022.

General discussion

Throughout this thesis, we studied how to measure individual inbreeding and its effects on individuals' phenotypes. We propose new methodological guidelines for researchers wanting to estimate inbreeding and inbreeding depression and we assess the inbreeding status of a large-scale wild metapopulation of barn owls. Inbreeding detection can be challenging, and we began by investigating which type of genomic data can be used with a common measure of inbreeding that relies on identical-by-descent (IBD) segments: runs of homozygosity (ROHs) [46, 43, 41, 63, 4]. We showed that the most popular software used to identify ROHs (PLINK [75]) can only be used if the genome has been sequenced at high-densities (namely with whole-genome sequencing and high-density SNP arrays). On the contrary, we found that a less popular method (RZOOROH [22]), can be applied at both high and low densities. Then, we compared the ability of different inbreeding coefficients (including but not limited to ROHs-based coefficients) to quantify the effects of inbreeding on phenotypic traits. We proposed a novel statistical method to take into account the non-independence of observations when quantifying inbreeding depression in populations with strong genetic structure. We showed that the quantification of inbreeding depression with non-independent data is greatly improved when a genetic relatedness matrix (GRM) based on allele-sharing is included in the (generalized linear) model. The most commonly used GRM, however, does not allow for an accurate estimation of the strength of inbreeding depression. After testing and comparing methods for both inbreeding and inbreeding depression quantification, we applied these methods to empirical data from the European barn owl (*Tyto alba*). We measured inbreeding and deleterious load in several barn owl populations of different sizes and looked for evidence of purging in island populations. We showed that island populations are more inbred than mainland populations and that this inbreeding is mostly due to ancient relatedness linked to islands' smaller population sizes. Additionally, we demonstrated that none of our populations has undergone purging of deleterious alleles. Finally, we quantified the effect of inbreeding on different traits in the Swiss barn owl population and find that inbred juveniles have a smaller bill on average than those who are outbred.

Why measuring inbreeding and inbreeding depression ?

In the first part of this thesis, we study how to accurately quantify inbreeding and its deleterious consequences and we show in chapters I and II that only the correct combination of data and methods can lead to reliable estimation of inbreeding and inbreeding depression. Several studies have shown that inbreeding is ubiquitous in nature [3, 37, 44, 19, 97] and how to accurately quantify it is an important topic and has been widely discussed in the literature [45, 41, 98, 42, 99, 6, 1]. As a result of climate change and the sixth biodiversity crisis [71, 12, 13, 11], a significant decrease in population sizes is occurring in nature which leads to an increase in inbreeding. Thus, it is essential to accurately quantify diversity and inbreeding in natural populations (as we do for the barn owl in chapters III and IV). It is crucial to identify endangered populations efficiently and develop appropriate conservation guidelines [27, 74]. Aside from being important for conservation, understanding inbreeding also plays an important role in human health since it has been linked to fertility reduction [72, 15], an increase in several diseases [49, 10], and a reduction in quantitative traits related to stature [10, 62] and cognition [10]. Finally, inbreeding is particularly prevalent among domestic species such as chicken [66, 50], cattle [64, 34, 77, 53], sheep [76, 60, 83], pigs [82, 37, 8] and aquacultures [55]. Besides feeding human populations, these species are also of great economic significance. Therefore, it is important to be able to manage their inbreeding load accurately in order to avoid population crash [59]. In fact, most domestic breeds today follow strict breeding programs in order to minimize inbreeding and its negative effects.

When studying the inbreeding status of populations, ROHs are particularly useful because they provide both an estimate of an individual inbreeding coefficient and information regarding the origin of this inbreeding via their distributions [10, 46]. For instance, Mastrangelo *et al.* (2016) [61] compared ROHs distribution from three dairy cattle breeds and found that Italian Holstein individuals harbored a high number of short ROHs. This suggests that inbreeding in this breed is mostly caused by ancient relatedness. On the other hand, individuals from the *Modicana* and *Cinisara* breeds harbored a high number of large ROHs, suggesting recent mating events between relatives. The authors concluded that a monitored breeding program for reducing consanguinity in these local breeds was necessary. It is therefore crucial that ROHs distributions are correctly estimated if they are used for taking decisions. We show in chapter I that not all methods can be used with reduced genomic representations and this

is particularly important because they are very popular due to their lower price, especially in conservation studies where funds are often limited.

Even though inbreeding can have deleterious consequences for individuals, it does not necessarily result in inbreeding depression or population extinction. In nature, populations with long-term small population sizes and high levels of inbreeding have been observed with little or no inbreeding depression [78, 2, 79]. Consequently, it is crucial not only to quantify inbreeding accurately, but also to determine whether a population (wild or domestic) suffers from inbreeding depression and requires active management. Indeed, we show in chapter II that inbreeding depression quantification can be strongly biased if the methods are not adapted. Identifying the type of management it requires is also important. It has been demonstrated that “simply” increasing the genetic diversity of a small potentially endangered population is not always the most effective method for population rescue and deep understanding of population history as well as the cause of inbreeding depression are needed [89, 36, 30].

The importance of methodology in biology

Methodological studies are critical, but they are often overlooked. Indeed, scientific research aims to understand biological processes from which interpretations and conclusions can be derived. The conclusions drawn from these studies may serve as a basis for making important decisions. We mentioned above that many studies aim at examining potential conservation issues and conservation statuses of populations and species to propose adapted conservation strategies [61, 74] and we showed in chapters I and II that accurate estimation of inbreeding and inbreeding depression can only be achieved with specific combinations and of data and methods. As another example, biomedical research plays a significant role in human health via the development and commercialization of new drugs [21]. The conclusions reached by scientists must, therefore, be accurate if they are to be used to make decisions. And to reach accurate conclusions, we need a robust methodology and a deep understanding of both data and methods. Moreover, it is imperative that these methods and models are rigorously validated through methodological studies as we did in chapter I and II. Especially in today’s world where technology is ever-evolving, it is crucial to ensure that we have reliable methods in place to deal with the large quantity and various quality of data being generated. For instance, inappropriate conservation management based on erroneous results produced by wrong combinations of data and methods can be detrimental to

biodiversity or, on the contrary, a waste of resources. Consequently, it is crucial that the research methodology, interpretations, and conclusions are accurate and carefully examined.

Implication for the barn owl

This thesis contributes to the growing body of literature concerning the characterization of inbreeding as well as the detection of inbreeding depression in the wild [86, 40, 43, 35, 68]. Even though barn owls are well studied, to our knowledge, we are the first to conduct such a large-scale assessment of their inbreeding status. We demonstrate that the European barn owl, despite increased inbreeding in island populations, displays high genetic diversity and low levels of inbreeding. Therefore, we confirm that this species is not endangered, as assessed by the [International Union for conservation in Nature](#) (IUCN) and does not require conservation measures. Despite its low inbreeding levels, we still detected inbreeding depression. In this context, inbreeding depression is not a concern: it only affects a small number of the Swiss barn owls' population and will theoretically not result in the Swiss population extinction. This is an important consideration, as barn owls are keystone species as they help regulate populations of small rodents and can prevent overgrazing in ecosystems.

Future works and limitations

In this thesis, we studied many aspects related to inbreeding but there is a lot more that could be done with enough time and resources. In the first chapter, we tested two different approaches for identifying IBD segments. As I mentioned in the discussion, there is a third approach [5, 70] which we did not test, and it would be interesting to see how it handles reduced genomic representations. Furthermore, confirmation of our results with empirical data would be helpful, even though it is difficult to find individuals who have been sequenced twice at different densities. We could however use whole-genome sequencing and subsample reads to mimic other sequencing techniques.

We proposed a novel approach for quantifying inbreeding depression in data with strong genetic structure in the second chapter. Firstly, the way we simulated inbreeding depression could be more realistic: we could simulate genetic data and attribute fitness values to specific loci to estimate individual fitness. Furthermore, we could explore different types of structure and try to find a minimum F_{ST} value from which taking the

non-independence of data is necessary for quantifying inbreeding depression. This would give precise guidelines to researchers wanting to quantify inbreeding depression as to whether they need to account for structure or not.

In the third chapter, we examine the degree of inbreeding in different barn owl populations and determine whether they are enriched in deleterious alleles. Since we were unable to polarize our data, we assumed that the minor allele was the deleterious allele. Ideally, it would be appropriate to redo all analyses (starting with the inclusion of out-groups in the genotyping) to allow for the polarization of our genomic data. As our group recently began studying the world barn owl phylogeny, we may be able to polarize our data by using their data. It would also be interesting to estimate the number of lethal equivalents, which is an alternative measure of genetic load (which does not require phenotypic data), in different populations of owls and compare them with populations on continental and island continents. In addition, it would be interesting to determine whether IBD segments are enriched or depleted in deleterious variants, and whether the results are the same for islands and continental populations. Finally, the effective populations' sizes we estimated are not very realistic and, while we can still use the relative values for comparison, it would be better to have more accurate estimates, either based on a better mutation rate (which we could estimate based on the barn owl data presented in this thesis) or by using specialized software [23, 54, 90].

Finally, in the last chapter of this thesis, we look for evidence of inbreeding depression in the Swiss barn owl population. For fitness-related traits, such as egg hatching probability, we did not detect inbreeding depression. This may be due to the fact that our sample size is very small. As a matter of fact, we had to discard a large amount of data because many barn owl eggs were cross-fostered in the last decade and it was impossible to identify the original parents of each egg. For hatched eggs, this can be done by sequencing both the parents and the juvenile. However, it is impossible to do so for eggs that did not hatch. New observations are added to the database every season and by sequencing new parents and adding more recent observations, we should be able to re-run these analyses with more observations. The initial sample selection included owls for which most of the phenotypic traits were known (including color, which was not used in this thesis but that other members of our groups are actively studying). Consequently, we did not sequence juveniles that died too early (before their plumage coloration could be assessed). We intend to sequence these individuals and perform inbreeding depression

analyses on juvenile survival rates. We are also interested in modeling recapture probability (as a proxy for survival probability) according to individuals inbreeding coefficients. As a final point, we did not include any genetic relatedness matrix in our analysis. It could be interesting to see how including it (instead of the pedigree-based matrix) would influence heritability and inbreeding depression strength estimation. Since our data do not have a substantial structure and there is no difference between the simple linear model and the model including the pedigree-based relatedness matrix, we do not expect substantial change in inbreeding depression strength estimation. However, it could improve our heritability estimation by reducing the associated confidence intervals.

Conclusion

In conclusion, this thesis emphasizes the importance of using accurate data and methods. First, the type of data matters: we cannot use any inbreeding detection method with any type of data. When we have the correct data, we must then determine the most appropriate method to both quantify inbreeding and estimate its impact on individuals. Therefore, the software and the inbreeding coefficient we use are extremely important factors. After inbreeding has been accurately quantified, the next challenge is to estimate how it affects individuals' traits. The analysis for inbreeding depression estimation must also be adapted to the studied population: the best inbreeding coefficient will actually depend on the population size and whether we should use the GRM will depend on whether our population is structured. Using the barn owl (*Tyto alba*) as a model organism, we demonstrate that with proper methods and data one can study inbreeding status and origin, as well as quantify inbreeding depression in populations. We confirm that the barn owl is not endangered but that inbreeding depression can still be found in its populations. In conclusion, this thesis contributed to a better understanding of the quantification of inbreeding and inbreeding depression, which are crucial for several fields such as conservation assessment, domestic populations management and humans' health.

References

- [1] S.W. Alemu, N.K. Kadri, C. Harland, P. Faux, C. Charlier, A. Caballero, and T. Druet. An evaluation of inbreeding measures using a whole-genome sequenced cattle pedigree. *Heredity*, 126(3):410–423, 2021.
- [2] A. Benazzo, E. Trucchi, J.A. Cahill, P. Maisano Delsler, S. Mona, M. Fumagalli, L. Bunnefeld, L. Cornetti, S. Ghirotto, M. Girardi, L. Ometto, A. Panziera, O. Rota-Stabelli, E. Zanetti, A. Karamanlidis, C. Groff, L. Paule, L. Gentile, C. Vilà, S. Vicario, L. Boitani, L. Orlando, S. Fuselli, C. Vernesi, B. Shapiro, P. Ciucci, and G. Bertorelle. Survival and divergence in a small group: The extraordinary genomic history of the endangered Apennine brown bear stragglers. *Proceedings of the National Academy of Sciences*, 114(45), 2017.
- [3] A.H. Bittles and M.L. Black. Consanguinity, human evolution, and complex diseases. *Proceedings of the National Academy of Sciences*, 107(suppl 1):1779–1786, 2010.
- [4] D.W. Bjelland, K.A. Weigel, N. Vukasinovic, and J.D. Nkrumah. Evaluation of inbreeding depression in Holstein cattle using whole-genome SNP markers and alternative measures of genomic inbreeding. *Journal of Dairy Science*, 96(7):4697–4706, 2013.
- [5] K.W. Broman and J.L. Weber. Long homozygous chromosomal segments in reference families from the centre d’étude du polymorphisme humain. *The American Journal of Human Genetics*, 65(6):1493–1500, 1999.
- [6] A. Caballero, B. Villanueva, and T. Druet. On the estimation of inbreeding depression using different measures of inbreeding from molecular markers. *Evolutionary Applications*, 14(2):416–428, 2021.
- [7] H.M. Cardullo and L.E. Holt. Ability of Infants to Taste PTC: Its Application in Cases of Doubtful Paternity. *Experimental Biology and Medicine*, 76(3):589–592, 1951.
- [8] J. Casellas, N. Ibáñez-Escriche, L. Varona, J.P. Rosas, and J.L. Noguera. Inbreeding depression load for litter size in *Entrepelado* and *Retinto* Iberian pig varieties. *Journal of Animal Science*, 97(5):1979–1986, 2019.
- [9] F.C. Ceballos and G. Álvarez. Royal dynasties as human inbreeding laboratories: the Habsburgs. *Heredity*, 111(2):114–121, 2013.
- [10] F.C. Ceballos, P.K. Joshi, D.W. Clark, . Ramsay, and J.F. Wilson. Runs of homozygosity: windows into population history and trait architecture. *Nature Reviews Genetics*, 19(4):220–234, 2018.
- [11] G. Ceballos and P.R. Ehrlich. The misunderstood sixth mass extinction. *Science*, 360(6393):1080–1081, 2018.
- [12] G. Ceballos, P.R. Ehrlich, A.D. Barnosky, A. García, R.M. Pringle, and T.M. Palmer. Accelerated modern human-induced species losses: Entering the sixth mass extinction. *Science Advances*, 1(5):e1400253, 2015.
- [13] G. Ceballos, P.R. Ehrlich, and R. Dirzo. Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proceedings of the National Academy of Sciences*, 114(30), 2017.
- [14] D. Charlesworth and J.H. Willis. The genetics of inbreeding depression. *Nature Reviews Genetics*, 10(11):783–796, 2009.
- [15] D.W. Clark, Y. Okada, K.H.S. Moore, D. Mason, N. Pirastu, I. Gandin, et al. Associations of autozygosity with a broad range of human phenotypes. *Nature communications*, 10(1):4957, 2019
- [16] J.F. Crow. *An introduction to population genetics theory*. Scientific Publishers, 1970.
- [17] T. Cumer, A.P. Machado, G. Dumont, V. Bontzorlos, R. Ceccherelli, M. Charter, K. Dichmann, N. Kassinis, R. Lourenço, F. Manzia, H-D. Martens, L. Prévost, M. Rakovic, I. Roque, F. Siverio, A. Roulin, and J. Goudet. Landscape and Climatic Variations Shaped Secondary Contacts amid Barn Owls of the Western Palearctic. *Molecular Biology and Evolution*, 39(1):msab343, 2022.
- [18] C. Darwin. *The Effects of Cross and Self Fertilisation in the Vegetable Kingdom*. 1877.
- [19] R.A. De Boer, R. Vega-Trejo, A. Kotrschal, and J.L. Fitzpatrick. Meta-analytic evidence that animals rarely avoid inbreeding. *Nature Ecology & Evolution*, 5(7):949–964, 2021.

- [20] J. Del Hoyo, A. Elliott, and J. Sargatal. *Handbook of the birds of the world*, volume 1. Lynx Ed., 1992.
- [21] J. Drews and S. Ryser. Drug Development: The role of innovation in drug development. *Nature Biotechnology*, 15(13):1318–1319, 1997.
- [22] T. Druet and M. Gautier. A model-based approach to characterize individual inbreeding at both global and local genomic scales. *Molecular Ecology*, 26(20):5820–5841, 2017.
- [23] A.J. Drummond, A. Rambaut, B. Shapiro, and O.G. Pybus. Bayesian Coalescent Inference of Past Population Dynamics from Molecular Sequences. *Molecular Biology and Evolution*, 22(5):1185–1192, 2005.
- [24] L. Duntsch, A. Whibley, P. de Villemereuil, P. Brekke, S. Bailey, J.G. Ewen, and A.W. Santure. Genomic signatures of inbreeding depression for a threatened Aotearoa New Zealand passerine. *Molecular Ecology*, 32(8):1893–1907, 2023.
- [25] N. Dussex, H.E. Morales, C. Grossen, L. Dalén, and C. Van Oosterhout. Purging and accumulation of genetic load in conservation. *Trends in Ecology & Evolution*, 38(10):961–969, 2023.
- [26] M.D.B. Eldridge, J.M. King, A.K. Loupis, P.B.S. Spencer, A.C. Taylor, L.C. Pope, and G.P. Hall. Unprecedented Low Levels of Genetic Variation and Inbreeding Depression in an Island Population of the Black-Footed Rock-Wallaby. *Conservation Biology*, 13(3):531–541, 1999.
- [27] S.P. Flanagan, B.R. Forester, E.K. Latch, S.N. Aitken, and S. Hoban. Guidelines for planning genomic assessment and monitoring of locally adaptive variation to inform species conservation. *Evolutionary Applications*, 11(7):1035–1052, 2018.
- [28] R. Frankham. Conservation genetics. *Annual Review of Genetics*, 29(1):305–327, 1995.
- [29] E. Furlan, J. Stoklosa, J. Griffiths, N. Gust, R. Ellis, R.M. Huggins, and A.R. Weeks. Small population size and extremely low levels of genetic diversity in island populations of the platypus, *Ornithorhynchus anatinus*. *Ecology and Evolution*, 2(4):844–857, 2012.
- [30] A. García-Dorado. Understanding and Predicting the Fitness Decline of Shrunk Populations: Inbreeding, Purging, Mutation, and Standard Selection. *Genetics*, 190(4):1461–1476, 2012.
- [31] S. Glémin. How are deleterious mutations purged ? Drift versus non-random mating. *Evolution*, 57(12):2678–2687, 2003.
- [32] C.G. Clark and C.G.C. Dickson. The life-history of *precis octavia* (cram.) (lepidoptera: *Nymphalidae*). *Journal of the Entomological Society of Southern Africa*, 20(2):257–259, 1957.
- [33] S-W. Guo. Variation in Genetic Identity among Relatives. *Human Heredity*, 46(2):61–70, 1996.
- [34] M.A. Gutiérrez-Reinoso, P.M. Aponte, and M. García-Herreros. A review of inbreeding depression in dairy cattle: current status, emerging control strategies, and future prospects. *Journal of Dairy Research*, 89(1):3–12, 2022.
- [35] K.A. Harrisson, M.J.L. Magrath, J.D.L. Yen, A. Pavlova, N. Murray, B. Quin, P. Menkhorst, K.A. Miller, K. Cartwright, and P. Sunnucks. Lifetime Fitness Costs of Inbreeding and Being Inbred in a Critically Endangered Bird. *Current Biology*, 29(16):2711–2717.e4, 2019.
- [36] P.W. Hedrick, R.O. Peterson, L.M. Vucetich, J.R. Adams, and J.A. Vucetich. Genetic rescue in isle royale wolves: genetic analysis and the collapse of the population. *Conservation genetics*, 15:1111–1121, 2014.
- [37] J. Herrero-Medrano, H-J. Megens, M.A. Groenen, G. Ramis, M. Bosse, M. Pérez-Enciso, and R.P. Crooijmans. Conservation genomic analysis of domestic and wild pig populations from the Iberian Peninsula. *BMC Genetics*, 14(1):106, 2013.
- [38] W.G. Hill and B.S. Weir. Variation in actual relationship among descendants of inbred individuals. *Genetics Research*, 94(5):267–274, 2012.
- [39] D.P. Howrigan, M.A. Simonson, and M.C. Keller. Detecting autozygosity through runs of homozygosity: A comparison of three autozygosity detection algorithms. *BMC Genomics*, 12(1):460, 2011.
- [40] J. Huisman, L.E.B. Kruuk, P.A. Ellis, T. Clutton-Brock, and J.M. Pemberton. Inbreeding depression across the lifespan in a wild mammal population. *Proceedings of the National Academy of Sciences*, 113(13):3585–3590, 2016.
- [41] M. Kardos, G. Luikart, and F.W. Allendorf. Measuring individual inbreeding in the age of genomics:

- marker-based measures are better than pedigrees. *Heredity*, 115(1):63–72, 2015.
- [42] M. Kardos, P. Nietlisbach, and P.W. Hedrick. How should we compare different genomic estimates of the strength of inbreeding depression? *Proceedings of the National Academy of Sciences*, 115(11), 2018.
- [43] M. Kardos, Y. Zhang, K.M. Parsons, A. Yunga, H. Kang, X. Xu, X. Liu, C.O. Matkin, P. Zhang, E.J. Ward, M.B. Hanson, C. Emmons, M.J. Ford, G. Fan, and S.i Li. Inbreeding depression explains killer whale population dynamics. *Nature Ecology & Evolution*, 7(5):675–686, 2023.
- [44] L.F Keller and D.M. Waller. Inbreeding effects in wild populations. *Trends in Ecology & Evolution*, 17(5):230–241, 2002.
- [45] M.C. Keller, P.M. Visscher, and M.E. Goddard. Quantification of Inbreeding Due to Distant Ancestors and Its Detection Using Dense Single Nucleotide Polymorphism Data. *Genetics*, 189(1):237–249, 2011.
- [46] M. Kirin, R. McQuillan, C.S. Franklin, H. Campbell, P.M. McKeigue, and J.F. Wilson. Genomic Runs of Homozygosity Record Population History and Consanguinity. *PLoS ONE*, 5(11):e13996, 2010.
- [47] M. Kirkpatrick and P. Jarne. The effects of a bottleneck on inbreeding depression and the genetic load. *The American Naturalist*, 155(2):154–167, 2000.
- [48] C. König and F. Weick. *Owls of the world*. A&C Black, 2008.
- [49] C.S. Ku, N. Naidoo, S.M. Teo, and Y. Pawitan. Regions of homozygosity and their impact on complex diseases and traits. *Human Genetics*, 129(1):1–15, 2011.
- [50] J.M. Larivière, J. Detilleux, P. Leroy, *et al.* Estimates of inbreeding rates in forty traditional belgian chicken breeds populations. *Archiv für Geflügelkunde*, 75(1):1–6, 2011.
- [51] E. Lavanchy and J. Goudet. Effect of reduced genomic representation on using runs of homozygosity for inbreeding characterization. *Molecular Ecology Resources*, 23(4):787–802, 2023.
- [52] G. Leroy. Inbreeding depression in livestock species: review and meta-analysis. *Animal Genetics*, 45(5):618–628, 2014.
- [53] G. Leroy, T. Mary-Huard, E. Verrier, S. Danvy, E. Charvolin, and C. Danchin-Burge. Methods to estimate effective population size using pedigree data: Examples in dog, sheep, cattle and horse. *Genetics Selection Evolution*, 45(1):1–10, 2013.
- [54] H. Li and R. Durbin. Inference of human population history from individual whole-genome sequences. *Nature*, 475(7357):493–496, 2011.
- [55] D.Y. Ma, H.L. Hu, and J. Kong. Inbreeding and its impact on aquaculture. *Journal of Fisheries of China*, 29(6):849–856, 2005.
- [56] A.P. Machado, T. Cumer, C. Iseli, E. Beaudoin, A-L. Ducrest, M. Dupasquier, N. Guex, K. Dichmann, R. Lourenço, J. Lusby, H-D. Martens, L. Prévost, D. Ramsden, A. Roulin, and J. Goudet. Unexpected post-glacial colonisation route explains the white colour of barn owls (*tyto alba*) from the british isles. *Molecular Ecology*, 31(2):482–497, 2021.
- [57] A.P. Machado, A. Topaloudis, T. Cumer, E. Lavanchy, V. Bontzorlos, R. Ceccherelli, M. Charter, N. Kassinis, P. Lymberakis, F. Manzia, A-L. Ducrest, M. Dupasquier, N. Guex, A. Roulin, and J. Goudet. Genomic consequences of colonisation, migration and genetic drift in barn owl insular populations of the eastern Mediterranean. *Molecular Ecology*, 31(5):1375–1388, 2022.
- [58] G. Malécot. *Les mathématiques de l'hérédité*. Masson & Cie Paris, 1948.
- [59] Food and Agriculture Organization (FAO). Building on gender, agrobiodiversity and local knowledge. A training manual. 2005.
- [60] S. Mastrangelo, E. Ciani, M.T. Sardina, G. Sottile, F. Pilla, B. Portolano, and BiOvIta Consortium. Runs of homozygosity reveal genome-wide autozygosity in Italian sheep breeds. *Animal Genetics*, 49(1):71–81, 2018.
- [61] S. Mastrangelo, M. Tolone, R. Di Gerlando, L. Fontanesi, M.T. Sardina, and B. Portolano. Genomic inbreeding estimation in small populations: evaluation of runs of homozygosity in three local dairy cattle breeds. *animal*, 10(05):746–754, 2016.
- [62] R. McQuillan, N. Eklund, N. Pirastu, M. Kuningas, B.P. McEvoy, T. Esko, T. Corre, G. Davies, M.

- Kaakinen, L-P. Lyytikäinen, K. Kristiansson, A.S. Havulinna, M. Gögele, V. Vitart, A. Tenesa, Y. Aulchenko, C. Hayward, A. Johansson, M. Boban, S. Ulivi, A. Robino, V. Boraska, W. Igl, S.H. Wild, L. Zgaga, N. Amin, E. Theodoratou, O. Polasek, G. Girotto, L.M. Lopez, C. Sala, J. Lahti, T. Laatikainen, I. Prokopenko, M. Kals, J. Viikari, J. Yang, A. Pouta, K. Estrada, A. Hofman, N. Freimer, N.G. Martin, M. Kähönen, L. Milani, M. Heliövaara, E. Vartiainen, K. Räikkönen, C. Masciullo, J.M. Starr, A.A. Hicks, L. Esposito, I. Kolcic, S.M. Farrington, B. Oostra, T. Zemunik, H. Campbell, M. Kirin, M. Pehlic, F. Faletta, D. Porteous, G. Pistis, E. Widén, V. Salomaa, S. Koskinen, K. Fischer, T. Lehtimäki, A. Heath, M.I. McCarthy, F. Rivadeneira, G.W. Montgomery, H. Tiemeier, A-L. Hartikainen, P.A.F. Madden, P. d’Adamo, N.D. Hastie, U. Gyllensten, A.F. Wright, C. M. van Duijn, M. Dunlop, I. Rudan, P. Gasparini, P.P. Pramstaller, I.J. Deary, D. Toniolo, J.G. Eriksson, A. Jula, O.T. Raitakari, A. Metspalu, M. Perola, M-R. Järvelin, A. Uitterlinden, P.M. Visscher, J.F. Wilson on behalf of the ROHgen Consortium. Evidence of inbreeding depression on human height. *PLOS Genetics*, 8(7):1–14, 2012.
- [63] R. McQuillan, A-L. Leutenegger, R. Abdel-Rahman, C.S. Franklin, M. Pericic, L. Barac-Lauc, N. Smolej-Narancic, B. Janicijevic, O. Polasek, A. Tenesa, A. K. MacLeod, S.M. Farrington, P. Rudan, C. Hayward, V. Vitart, I. Rudan, S.H. Wild, M.G. Dunlop, A.F. Wright, H. Campbell, and J.F. Wilson. Runs of Homozygosity in European Populations. *The American Journal of Human Genetics*, 83(3):359–372, 2008.
- [64] F. Miglior, B. Szkotnicki, and E.B. Burnside. Analysis of Levels of Inbreeding and Inbreeding Depression in Jersey Cattle. *Journal of Dairy Science*, 75(4):1112–1118, 1992.
- [65] C. Mora, D.P. Tittensor, S. Adl, A.G. B. Simpson, and B. Worm. How Many Species Are There on Earth and in the Ocean? *PLoS Biology*, 9(8):e1001127, 2011.
- [66] B. Mtileni, K. Dzama, K. Nephawe, and C. Rhode. Estimates of effective population size and inbreeding in South African indigenous chicken populations: implications for the conservation of unique genetic resources. *Tropical Animal Health and Production*, 48(5):943–950, 2016.
- [67] I. Newton. *Population ecology of raptors*. A&C Black, 2010.
- [68] A.K. Niskanen, A.M. Billing, H. Holand, I.J. Hagen, Y.G. Araya-Ajoy, A. Husby, B. Rønning, A.M. Myhre, P.S. Ranke, T. Kvalnes, H. Pärn, T.H. Ringsby, S. Lien, B-E. Sæther, S. Muff, and H. Jensen. Consistent scaling of inbreeding depression in space and time in a house sparrow metapopulation. *Proceedings of the National Academy of Sciences*, 117(25):14584–14592, 2020.
- [69] T.J. Pemberton, D. Absher, M.W. Feldman, R.M. Myers, N.A. Rosenberg, and J.Z. Li. Genomic Patterns of Homozygosity in Worldwide Human Populations. *The American Journal of Human Genetics*, 91(2):275–292, 2012.
- [70] T.J. Pemberton, D. Absher, M.W. Feldman, R.M. Myers, N.A. Rosenberg, and J.Z. Li. Genomic patterns of homozygosity in worldwide human populations. *The American Journal of Human Genetics*, 91(2):275–292, 2012.
- [71] S.L. Pimm, C.N. Jenkins, R. Abell, T.M. Brooks, J.L. Gittleman, L.N. Joppa, P.H. Raven, C.M. Roberts, and J.O. Sexton. The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, 344(6187):1246752, 2014.
- [72] E. Postma, L. Martini, and P. Martini. Inbred women in a small and isolated Swiss village have fewer children. *Journal of Evolutionary Biology*, 23(7):1468–1474, 2010.
- [73] J.K. Pritchard. Are Rare Variants Responsible for Susceptibility to Complex Diseases? *The American Journal of Human Genetics*, 69(1):124–137, 2001.
- [74] V. Pumpitakkul, W. Chetruengchai, C. Srichomthong, C. Phokaew, W. Pootakham, C. Sonthirod, W. Nawae, S. Tongsimma, P. Wangkumhang, A. Wilantho, Y. Utara, A. Thongpakdee, S. Sanannu, U. Maikaew, S. Khuntawee, W. Changpetch, P. Phromwat, K. Raschasin, P. Sarnkhaeveerakul, P. Supapannachart, W. Buthasane, B.S. Pukazhenthii, K-P. Koepfli, P. Suriyaphol, S. Tangphatsornruang, G. Suriyaphol, and V. Shotelersuk. Comparative genomics and genome-wide SNPs of endangered Eld’s deer provide breeder selection for inbreeding avoidance. *Scientific Reports*, 13(1):19806, 2023.
- [75] S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M.A.R. Ferreira, D. Bender, J. Maller, Pa. Sklar, P.I.W. de Bakker, M.J. Daly, and P.C. Sham. *PLINK*: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, 81(3):559–575,

2007.

- [76] D.C. Purfield, S. McParland, E. Wall, and D.P. Berry. The distribution of runs of homozygosity and selection signatures in six commercial meat sheep breeds. *PLOS ONE*, 12(5):e0176780, 2017.
- [77] A. Robertson. Inbreeding in artificial selection programmes. *Genetical Research*, 2(2):189–194, 1961.
- [78] J.A. Robinson, C. Brown, B.Y. Kim, K.E. Lohmueller, and R.K. Wayne. Purging of Strongly Deleterious Mutations Explains Long-Term Persistence and Absence of Inbreeding Depression in Island Foxes. *Current Biology*, 28(21):3487–3494.e4, 2018.
- [79] J.A. Robinson, C.C. Kyriazis, S.F. Nigenda-Morales, A.C. Beichman, L. Rojas-Bracho, K.M. Robertson, M.C. Fontaine, R.K. Wayne, K.E. Lohmueller, B.L. Taylor, and P.A. Morin. The critically endangered vaquita is not doomed to extinction by inbreeding depression. *Science*, 376(6593):635–639, 2022.
- [80] A. Romano, R. Séchaud, and A. Roulin. Global biogeographical patterns in the diet of a cosmopolitan avian predator. *Journal of Biogeography*, 47(7):1467–1481, 2020.
- [81] A. Roulin. *Tyto alba* barn owl. *BWP Update*, 4(2):115–138, 2002.
- [82] G. Schiavo, S. Bovo, F. Bertolini, S. Tinarelli, S. Dall’Olio, L. Nanni Costa, M. Gallo, and L. Fontanesi. Comparative evaluation of genomic inbreeding parameters in seven commercial and autochthonous pig breeds. *animal*, 14(5):910–920, 2020.
- [83] H. Signer-Hasler, A. Burren, P. Ammann, C. Droege Mueller, and C. Flury. Extent of genomic inbreeding in Swiss sheep and goat breeds. *Agrarforschung Schweiz*, 10(10):372–379, 2019.
- [84] D.W. Snow and S. Cramp. *The Complete Birds of the Western Palearctic*. Oxford University Press, 1998.
- [85] D. Speed and D.J. Balding. Relatedness in the post-genomic era: is it still useful? *Nature Reviews Genetics*, 16(1):33–44, 2015.
- [86] M.A. Stoffel, S.E. Johnston, J.G. Pilkington, and J.M. Pemberton. Genetic architecture and lifetime dynamics of inbreeding depression in a wild mammal. *Nature Communications*, 12(1):2972, 2021.
- [87] N.A. Swinford, S.P. Prall, S. Gopalan, C.M. Williams, J. Sheehama, B.A. Scelza, and B.M. Henn. Increased homozygosity due to endogamy results in fitness consequences in a human population. *Proceedings of the National Academy of Sciences*, 120(43):e2309552120, 2023.
- [88] I. Taylor. *Barn owls: predator-prey relationships and conservation*. Cambridge University Press, 2003.
- [89] J.C. Teixeira and C.D. Huber. The inflated significance of neutral genetic diversity in conservation genetics. *Proceedings of the National Academy of Sciences*, 118(10):e2015096118, 2021.
- [90] J. Terhorst, J.A. Kamm, and Y.S. Song. Robust and scalable inference of population history from hundreds of unphased whole genomes. *Nature Genetics*, 49(2):303–309, 2017.
- [91] E.A. Thompson. Identity by Descent: Variation in Meiosis, Across Genomes, and in Populations. *Genetics*, 194(2):301–326, 2013.
- [92] V. Uva, M. Päckert, A. Cibois, L. Fumagalli, and A. Roulin. Comprehensive molecular phylogeny of barn owls and relatives (Family: *Tytonidae*), and their six major Pleistocene radiations. *Molecular Phylogenetics and Evolution*, 125:127–137, 2018.
- [93] B.S. Weir and J. Goudet. A Unified Characterization of Population Structure and Relatedness. *Genetics*, 206(4):2085–2103, 2017.
- [94] J.J. Wiens. How many species are there on Earth? Progress and problems. *PLOS Biology*, 21(11):e3002388, 2023.
- [95] S. Wright. Coefficients of Inbreeding and Relationship. *The American Naturalist*, 56(645):330–338, 1922.
- [96] J. Yang, S. Lee, M.E. Goddard, and P.M. Visscher. *GCTA*: A Tool for Genome-wide Complex Trait Analysis. *The American Journal of Human Genetics*, 88(1):76–82, 2011.
- [97] L. Yengo, N.R. Wray, and P.M. Visscher. Extreme inbreeding in a European ancestry sample from the contemporary UK population. *Nature Communications*, 10(1):3719, 2019.
- [98] L. Yengo, Z. Zhu, N.R. Wray, B.S. Weir, J. Yang, M.R. Robinson, and P.M. Visscher. Detection and quantification of inbreeding depression for complex traits from SNP data. *Proceedings of the National Academy of Sciences*, 114(32):8602–8607, 2017.

[99]Q.S. Zhang, J. Goudet, and B.S. Weir. Rank-invariant estimation of inbreeding coefficients. *Heredity*, 128(1):1–10, 2022.

MOLECULAR ECOLOGY RESOURCES

Supplemental Information for:

**Effect of reduced genomic representation on using runs of
homozygosity for inbreeding characterization**

Eléonore Lavanchy & Jérôme Goudet

Table of Contents:

Simulations: small and large populations	Page 2
Simulations: cattle population	Page 3
ROHs calling with <i>PLINK</i>: parameters optimization	Page 5
Supplementary results	Page 10
Table S3	Page 23
References	Page 25

MOLECULAR ECOLOGY

RESOURCES

SUPPLEMENTARY MATERIAL & METHODS

Simulations: small and large populations

The genetic map was simulated with *FREGENE* (Chadeau-Hyam et al., 2008) with default parameters to mimic a human-like recombination map. Input files with default parameters can be accessed while downloading the *FREGENE* software in *Fregene/Example/data/*. Only the `<CHROMO_LENGTH>` parameter was modified to simulate chromosomes with size of 100KB.

We performed two rounds of simulations with *SLiM3* (Haller & Messer, 2019) to force all chromosomes from the same simulation replicate to have the same history/pedigree. For each simulation replicate, the aim of the first round was to generate the pedigree which was then “applied” 30 times (for 30 chromosomes) in the second round of simulations. In the first round of simulations, we used a non-Wright-Fisher model with overlapping generations: individuals ages varied between 0 and 3 years old. Ages were uniformly assigned to individuals at the first generation. The population size was regulated at each generation at the end of the simulation cycle: only N_e individuals (i.e. 1,000 for the small and 10,000 for the large population) with the higher relative fitness survived through the next generation. In addition to the density-dependency fitness, individual fitness was relative to age: probability to survive to the next generation was 0.9 for 0-year old individuals, 0.7 for 1-year old individuals, 0.5 for 2-years old individuals and 0 for 3-years old individuals. A reproduction callback (which defines

MOLECULAR ECOLOGY

RESOURCES

reproduction events in *SLiM3*) was called at each generation: N_e individuals (i.e. 1,000 for the small and 10,000 for the large population) were sampled for reproduction (without replacement). Individual probabilities of being selected as a mate were age-dependent: 0 for 0 years old individuals, 0 for 1-year old individuals, 0.6 for 2-years old individuals and 0.4 for 3-years old individuals. For each of these N_e individuals, a mate was then selected among the other adults (i.e. older than 1-year old) of the population, creating one new individual in the population. For each potential mate, the probability of being chosen as a mate was equal to 10 times the relatedness (estimated with *SLiM3*) between the potential mate and the focal individual + 0.01 multiplied by 0.6 for 2-years old individuals and 0.4 for 3-years old individuals. The first part was to ensure that some inbreeding occurred at each generation. Mating and death events were recorded for generation. During the second round of simulations, these mating and deaths events were applied to the 30 chromosomes (each with a different genetic map).

The burn-in were performed with *python* via *recapitation* in *msprime* (Kelleher et al., 2016), with a constant recombination rate of 1×10^{-8} . Mutations were added at the end of the simulation based on a human-like mutation rate of 2.5×10^{-8} .

Simulations: cattle population

The genetic map from male Holstein cattle was obtained from a study from Qanbari & Wittenburg (2020). Chromosome sizes correspond to the true chromosome sizes from *Bos Taurus*.

MOLECULAR ECOLOGY

RESOURCES

The history of the simulation was inspired by a paper from Frantz et al. (2020). For this population, we first simulated the burn-in with *msprime*: We simulated two populations: a wild population with 50,000 individuals and a domestic population with 1,500 individuals which split 10,000 generations ago. The migration rate between both populations was constant and equal to 3×10^{-5} . During this burn-in, the recombination rate was constant and equal to 1×10^{-8} . At the end of the burn-in, 200 individuals were randomly sampled to mimic the strong artificial pressure applied 200 generations ago during many current breeds creation. Individual metadata were modified to fit a non-Wright-Fisher model: half of the individual were assigned a female gender and the other half a male gender. Individual ages were randomly assigned between 0 and 3 years old.

The last 200 generations were performed in *SLiM3*. Since our real pedigree only covered 57 years maximum, we performed a first round of simulation and recorded all mating events. We then used this simulated pedigree to complete the real one: each founder of the real pedigree (i.e. each individual from which parents are unknown) was randomly linked to a simulated individual from the same generation, consequently receiving his genealogy (from the simulated pedigree). In other words, we used the simulated genealogies to complete the real ones. Afterwards, we trimmed this completed pedigree to keep only the individuals and mating events resulting in the individuals from the last generation of the real pedigree. Then, we ran a second round of simulations where we simply applied this new complete pedigree to all chromosome of the cattle genome in *SLiM3*. Finally, we used *msprime* to add all mutations with a mutation rate of 2.5×10^{-8} .

MOLECULAR ECOLOGY

RESOURCES

and to subsample the individuals from the last generation only. This corresponds to our WGS data for the cattle population.

ROHs calling with *PLINK*: parameters optimization

The *PLINK* parameters used for ROHs calling varied according to the SNP density in the dataset. Hereafter are the list of the different parameters and how they were adjusted:

- Window size (*--homozyg-window-snp*): the window size (in number of SNPs) was positively correlated with SNPs density: we used larger windows for higher SNP densities. More precisely in reduced densities datasets the window-size was chosen to correspond, on average, to a segment with length corresponding to our minimum size threshold: 100KB as proposed by Dutsch et al., (2021): $Window\ size = SNP\ density \left[\frac{\#SNPs}{KB} \right] * 100$. However, we set a maximum threshold of 100 and a minimum threshold of 30 because using no minimum value resulted in low correlations between $F_{ROH\ RAD}$ and F_{IBD} (FIG S9). For WGS, we simply used 100.
- Minimum SNP density (*--homozyg-density*): the minimum SNP density [per KB] required for a homozygous segment to be called a ROH. This parameter also positively correlated with SNP density: a smaller density was required for dataset with low densities. Concerning RAD-sequencing and both arrays, the minimum density was set to 50KB. This means that at least one SNP must be present every 50KB. Concerning WGS,

MOLECULAR ECOLOGY

RESOURCES

we decided to include ROHs with at least 1 SNP for 100KB (from Purfield et al., (2012, 2017); Stoffel et al., (2021) cited in Duntsch et al., (2021)).

- Minimum number of SNPs per ROH (*--homozyg-snp*): we used the same value to what was set for the *--homozyg-window-snp* parameter. For RAD sequencing and SNP arrays, this threshold was chosen based on the mean SNP density in the dataset: $\#SNPs = SNP\ density \left[\frac{\#SNPs}{KB} \right] * 100$ with a maximum value of 100 and a minimum value of 30. This value was equal to 100 for WGS.
- Maximum distance in between two SNPs for them to be in the same ROH (*--homozyg-gap*): this value was set to default parameter 1Mb for the all the datasets. Using a lower value resulted in a strong overestimation of small ROHs and an underestimation of large ROHs (FIG S10).
- Maximum number of heterozygous SNPs per ROH (*--homozyg-het*): Since we have simulated data, we don't need to account for sequencing errors, only for *de novo* mutations. For RAD-sequencing, we allowed for one heterozygous SNP per ROH. Concerning both SNP arrays, we allowed for 0 heterozygous SNPs since the SNPs have been filtered on MAF and should not contain any *de novo* mutations. Finally, we authorized 8 heterozygous SNPs per ROH with WGS because it resulted in the best ROHs distributions estimation. This number is fairly high but is due to the simulated

MOLECULAR ECOLOGY

RESOURCES

data: we have many very rare variants which are normally filtered out during the quality control step during SNP calling with real data.

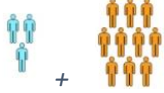
- Minimum scanning window hit rate (*--homozyg-window-treshold*): we decided to keep the default parameter 0.05 as seen in many other studies (Ceballos et al., 2018; Ferenčaković et al., 2013; Meyermans et al., 2020) cited in Duntsch et al., (2021).
- Maximum number of heterozygous SNPs per window (*--homozyg-window-het*): We authorized a maximum of two SNPs per window for RAD-sequencing and WGS and 0 for both SNP arrays (because SNPs have been filtered on $MAF > 0.05$).

ROHs calling with *PLINK*: parameters optimization; Cattle population

As mentioned in the main text, we had to increase the number of heterozygous SNPs allowed in a ROH (*--homozyg-het*) to 64 with WGS data in the cattle population. This is because the inbreeding is strong in this population and resulted in very long IBD segments. Consequently, higher numbers of *de novo* mutation were likely to occur in these IBD segments. However, we did not vary the number of heterozygous sites allowed in a window (*--homozyg-window-het*) to ensure that these heterozygous sites would not be too close to each other in the same ROH. We want to stress that we were able to perform parameter optimization for WGS with *PLINK* because we have access to the true fraction of genome within IBD segments which is not the case with empirical data.

MOLECULAR ECOLOGY RESOURCES

Table S1: PLINK parameters used for each sequencing technique for the small and large populations:



PARAMETERS	RAD sequencing	SNP arrays	WGS
Window size (<i>--homozyg-window-snp</i>)	$\left[\frac{\# \text{ SNPs}}{\text{KB}} \right] * 100$ Max = 100 Min = 30	$\left[\frac{\# \text{ SNPs}}{\text{KB}} \right] * 100$ Max = 100 Min = 30	100
Minimum SNP density (<i>--homozyg-density</i>)	50	50	100
Minimum number of SNPs per ROH (<i>--homozyg-snp</i>)	$\left[\frac{\# \text{ SNPs}}{\text{KB}} \right] * 100$ Max = 100 Min = 30	$\left[\frac{\# \text{ SNPs}}{\text{KB}} \right] * 100$ Max = 100 Min = 30	100
Maximum distance in between two SNPs for them to be in the same ROH (<i>--homozyg-gap</i>)	1000	1000	1000
Maximum number of heterozygous SNPs per ROH (<i>--homozyg-het</i>)	1	0	8
Minimum scanning window hit rate (<i>--homozyg-window-treshold</i>)	0.05	0.05	0.05
Maximum number of heterozygous SNPs per window (<i>--homozyg-window-het</i>)	2	0	2

MOLECULAR ECOLOGY RESOURCES

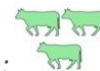


Table S2: PLINK parameters used for each sequencing technique for cattle population:

PARAMETERS	SNP arrays	WGS
Window size (<i>--homozyg-window-snp</i>)	$\left[\frac{\# \text{ SNPs}}{\text{KB}} \right] * 100$ Max = 100 Min = 30	100
Minimum SNP density (<i>--homozyg-density</i>)	50	100
Minimum number of SNPs per ROH (<i>--homozyg-snp</i>)	$\left[\frac{\# \text{ SNPs}}{\text{KB}} \right] * 100$ Max = 100 Min = 30	100
Maximum distance in between two SNPs for them to be in the same ROH (<i>--homozyg-gap</i>)	1000	1000
Maximum number of heterozygous SNPs per ROH (<i>--homozyg-het</i>)	0	64
Minimum scanning window hit rate (<i>--homozyg-window- treshold</i>)	0.05	0.05
Maximum number of heterozygous SNPs per window (<i>--homozyg-window-het</i>)	0	2

MOLECULAR ECOLOGY RESOURCES

SUPPLEMENTARY RESULTS

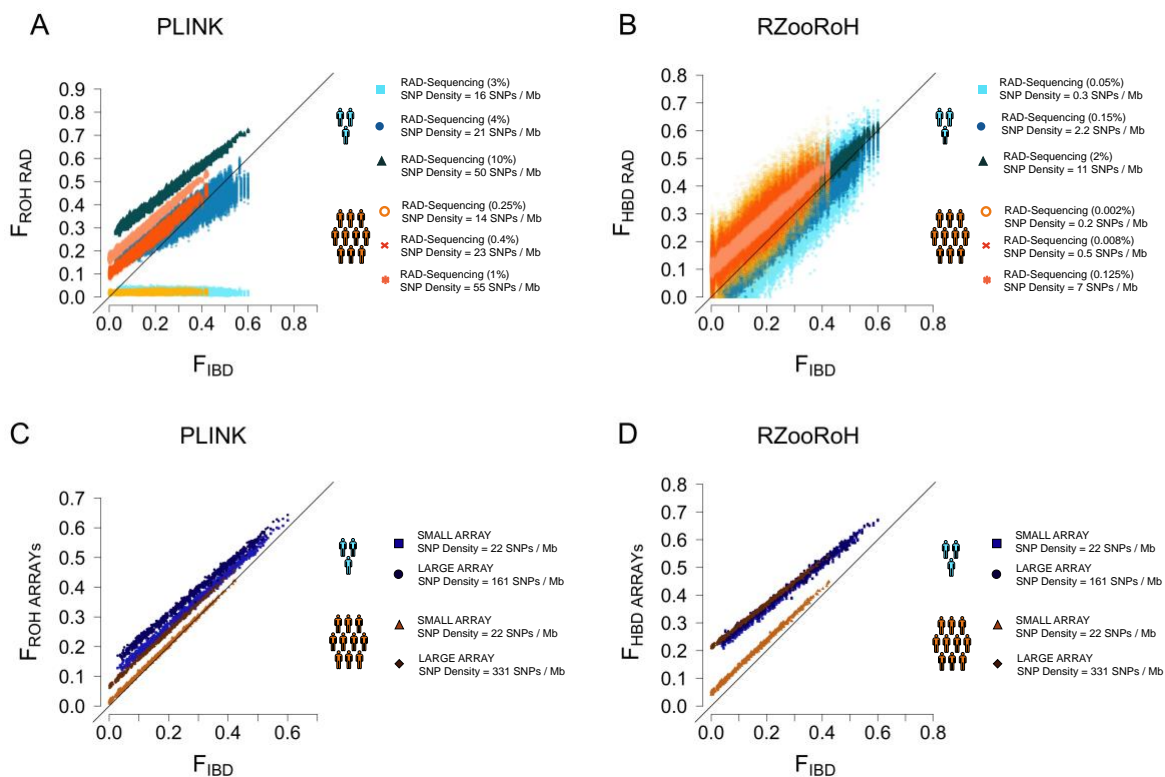


FIG S1: Comparison between F_{ROH} and F_{HBD} estimated with reduced genomic data on the y axis according to F_{IBD} (the true fraction of genome within IBD segments coalescing less than 100 reproductive cycles) on the x axis. Each point represents one individual (for one subsampling replicate within one simulation replicate). The black line represents the equality line ($x = y$). Blue points represent individuals from the small population and orangish from the large population. Within these two colors categories, a change in shade represents an increase in fraction of genome subsampled (indicated between the parentheses for RAD-sequencing). **A:** Subsampling was performed mimicking RAD-sequencing and ROHs were called with PLINK. **B:** Subsampling was performed mimicking RAD-sequencing and HBD segments were called with RZooRoH. **C:** Subsampling was performed mimicking SNPs arrays and ROHs were called with PLINK. **D:** Subsampling was performed mimicking SNPs arrays and HBD segments were called with RZooRoH.

MOLECULAR ECOLOGY RESOURCES

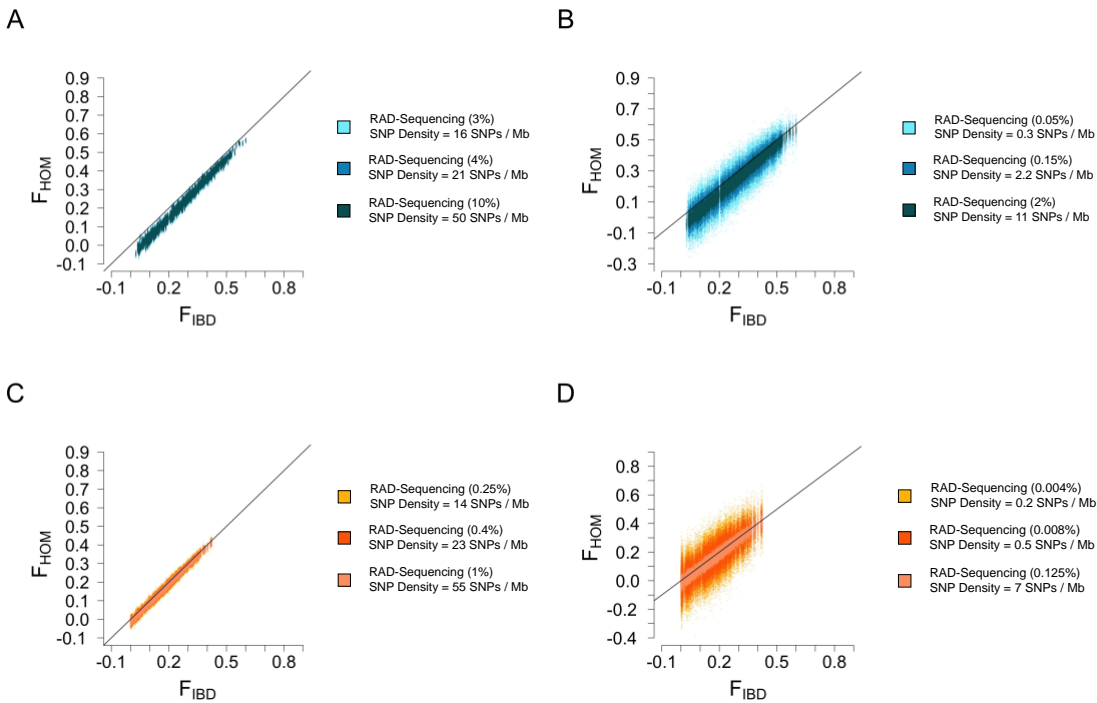


FIG S2: Comparison between F_{HOM} estimated with subsampled data mimicking RAD-sequencing (y axis) according to F_{IBD} , the true fraction of genome within IBD segments coalescing less than 100 reproductive cycles ago (x axis). **A:** Small population, fractions of genome presented are the same as the fractions of genomes used for ROHs analyses with PLINK (minimum size 100KB) and RAD-sequencing in FIG 4 and S1A; **B:** Small population, fractions of genome presented are the same as the fractions of genomes used for HBD segments analyses with RZooRoH and RAD-sequencing in FIG 4 and S1B; **C:** Large population, fractions of genome presented are the same as the fractions of genomes used for ROHs analyses with PLINK (minimum size 100KB) and RAD-sequencing in FIG 4 and S1A; **D:** Large population, fractions of genome presented are the same as the fractions of genomes used for HBD segments analyses with RZooRoH and RAD-sequencing in FIG 4 and S1B.

MOLECULAR ECOLOGY RESOURCES

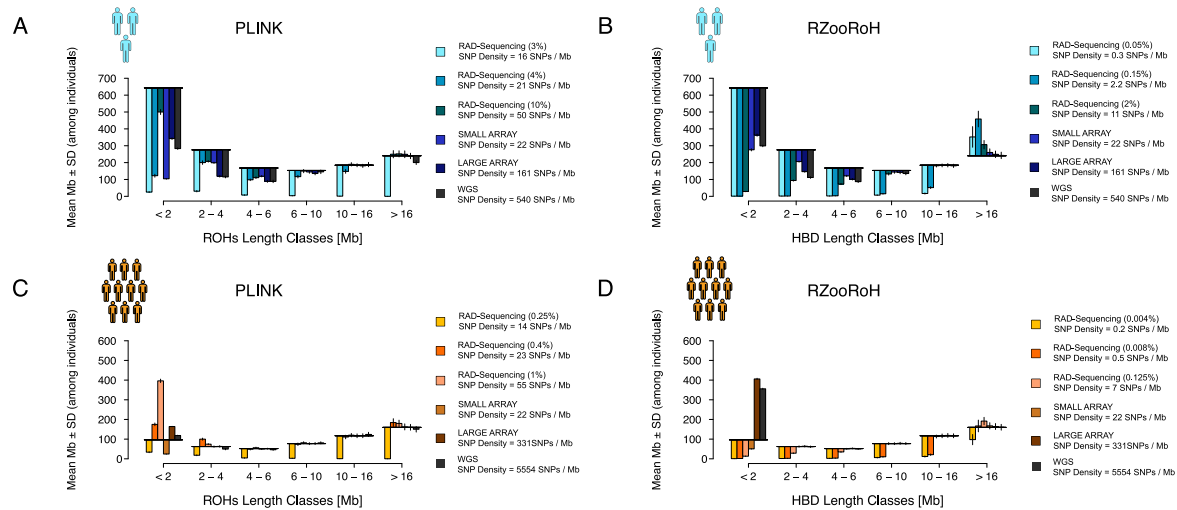


FIG S3: Comparison of ROHs (panels A and C) and HBD segments (panels B and D) distributions between the different sequencing methods and the true distributions of ‘older’ IBD segments (defined as segments which coalesced less than 1,000 reproductive cycles ago). Black horizontal lines correspond to total IBD segments lengths per individuals (y axis) falling into the different length classes (x axis). Barplots show the mean (\pm sd) difference between the mean total length of IBD segments and their estimation for each sequencing method. Barplots below the horizontal black line indicate an underestimation while barplots above the horizontal black line indicate an overestimation of the total length of segments. Mean (and sd) are among individuals, simulation and subsampling replicates. **A:** ROHs distributions from the small population; ROHs were called with PLINK. **B:** HBD segments distributions from the small population; HBD segments were called with RZooRoH. **C:** ROHs distributions from the large population; ROHs were called with PLINK. **D:** HBD segments distributions from the large population; HBD segments were called with RZooRoH.

MOLECULAR ECOLOGY RESOURCES

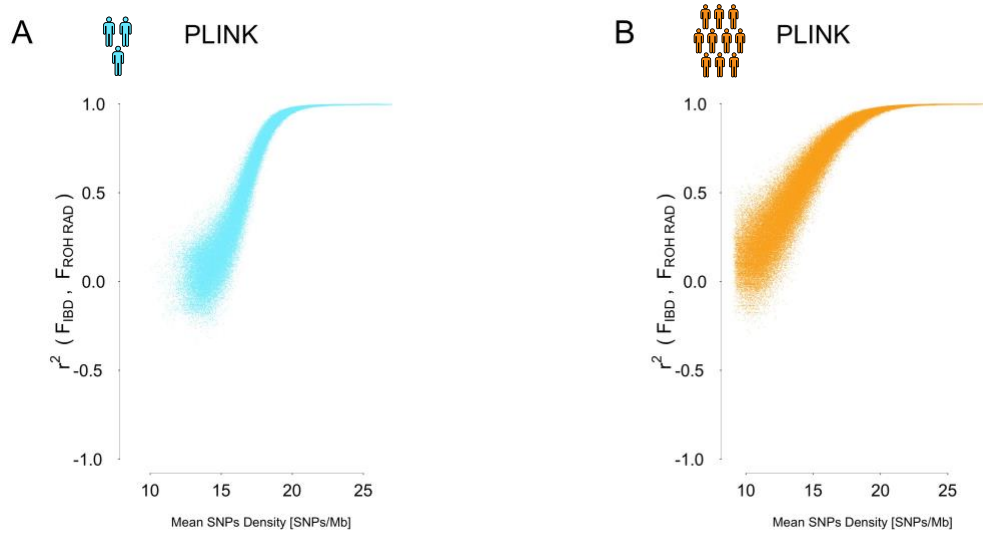


FIG S4: Correlation between F_{ROH} estimated with RAD-sequencing and F_{IBD} , the true fraction of genome within IBD segments (coalescing less than 100 reproductive cycles ago) according to SNP density in the reduced dataset for both the small and large population. ROHs were called with PLINK default parameters. **A:** Small population; **B:** Large population.

MOLECULAR ECOLOGY RESOURCES

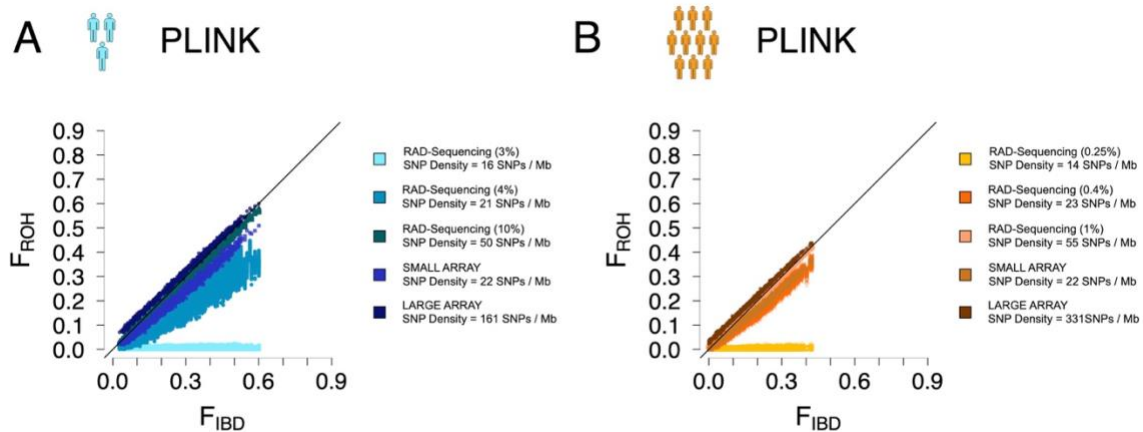


FIG S5: Comparison between F_{ROH} (called with PLINK default parameters) estimated with reduced genomic data on the y axis according to F_{IBD} (the true fraction of genome within IBD segments coalescing less than 100 reproductive cycles ago) on the x axis. Each point represents one individual (for one subsampling replicate within one simulation replicate). The black line represents the equality line ($x = y$). Blue points represent individuals from the small population and orangish from the large population. Within these two colors categories, a change in shade represents an increase in fraction of genome subsampled (indicated between the parentheses for RAD-sequencing). ROHs were called with PLINK default parameters. Fractions of genome presented are the same as the fractions of genomes used for ROHs analyses with PLINK in FIG 4, S1 and S2. **A:** Small population **B:** Large population.

MOLECULAR ECOLOGY RESOURCES

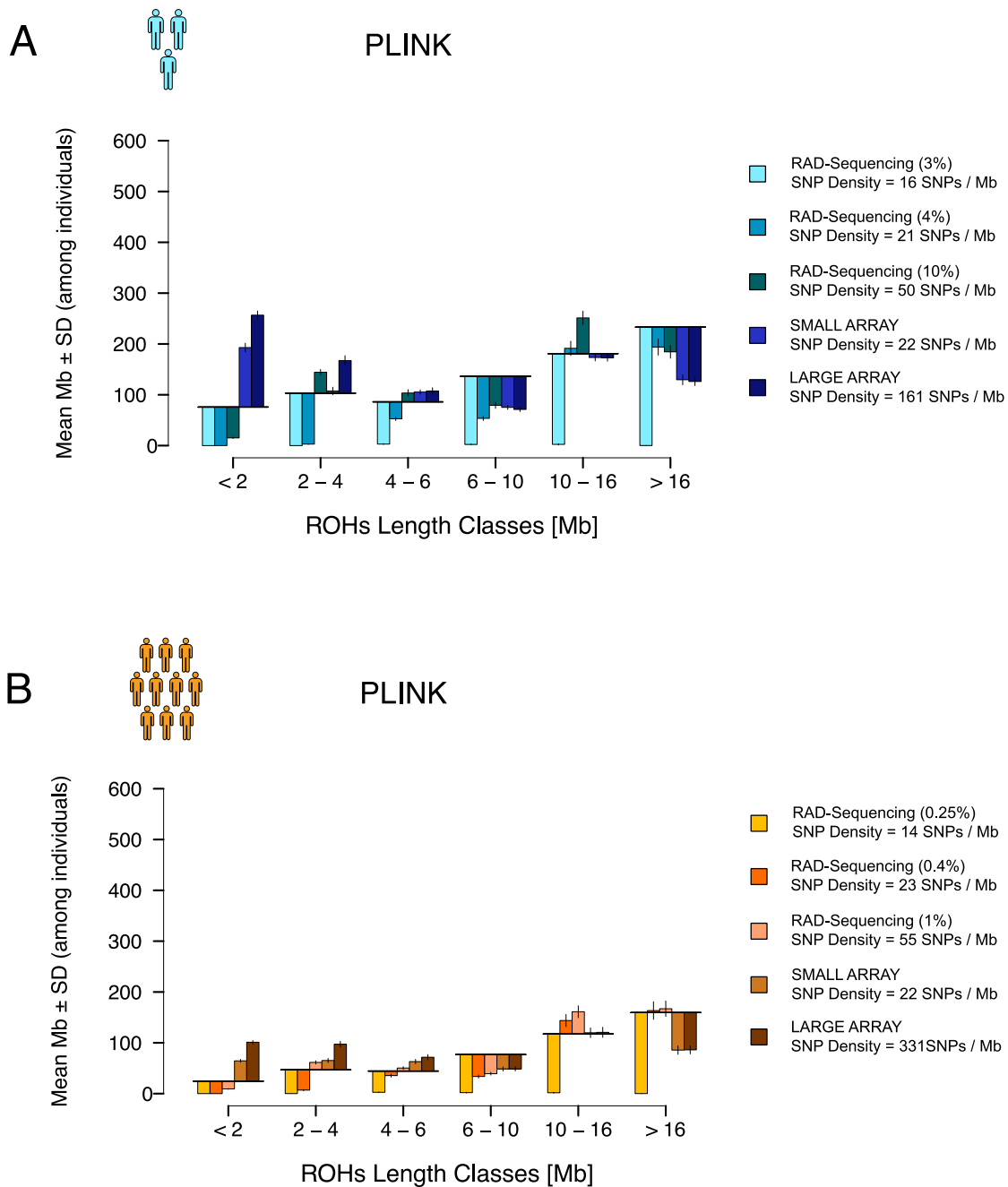


FIG S6: Comparison of ROHs (called with PLINK default parameters) distributions between the different sequencing methods and the true distributions of IBD segments (defined as segments which coalesced less than 100 reproductive cycles ago). Black horizontal lines correspond to total IBD segments lengths per individuals (y axis) falling into the different length classes (x axis). Barplots show

MOLECULAR ECOLOGY RESOURCES

*the mean (\pm sd) difference between the mean total length of IBD segment and their estimation for each sequencing method. Barplots below the horizontal black line indicate an underestimation while barplots above the horizontal black line indicate an overestimation of the total length of segments. Mean (and sd) are among individuals, simulation and subsampling replicates. ROHs were called with PLINK default parameters. Fractions of genome presented are the same as the fractions of genomes used for ROHs analyses with PLINK in FIG 4, S1, S2 and S5. **A:** Small population **B:** Large population.*

MOLECULAR ECOLOGY RESOURCES

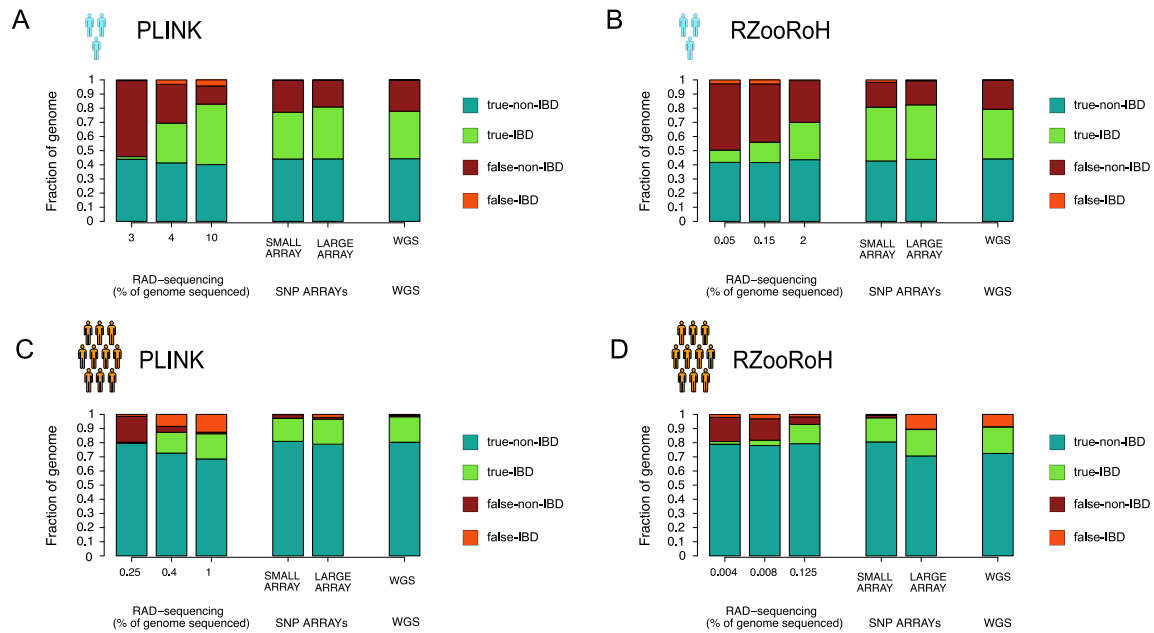


FIG S7: Segments are considered IBD if they coalesce less than 1,000 reproductive cycles ago. For each sequencing method: fraction of genome correctly assigned outside IBD segments (true-non-IBD), correctly assigned within IBD segments (true-IBD), incorrectly assigned outside IBD segments (false-non-IBD) and incorrectly assigned within IBD segments (false-IBD) are represented. Values are averaged among individuals as well as both simulation and subsampling replicates. **A:** ROHs distributions from the small population; ROHs were called with PLINK. **B:** HBD segments distributions from the small population; HBD segments were called with RZooRoH. **C:** ROHs distributions from the large population; ROHs were called with PLINK. **D:** HBD segments distributions from the large population; HBD segments were called with RZooRoH.

MOLECULAR ECOLOGY RESOURCES

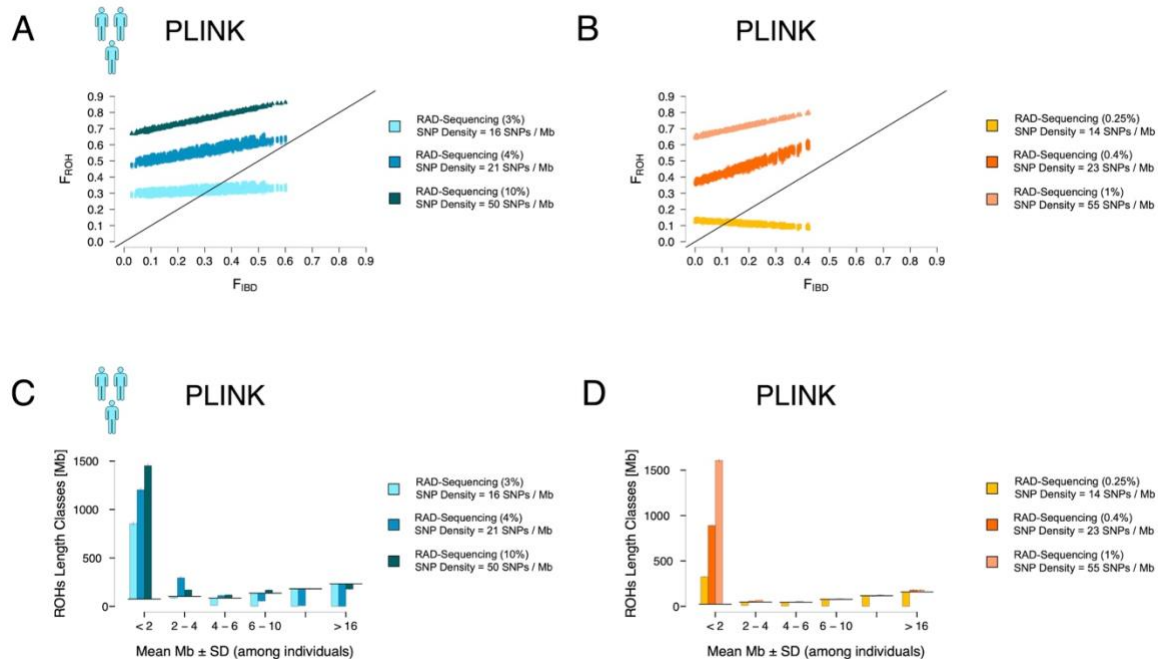


FIG S8: ROHs have been called with PLINK and a maximum length allowed between two adjacent SNPs (--homozyg-gap) of 1000. --homozyg-window-snp and --homozyg-snp were dependent on the SNP density in the reduced dataset with a no minimum value and a maximum value of 100. **A:** Comparison between F_{ROH} estimated with RAD-sequencing on the y axis according to F_{IBD} (the true fraction of genome within IBD segments coalescing less than 100 reproductive cycles ago) on the x axis. Each point represents one individual (for one subsampling replicate within one simulation replicate). The black line represents the equality line ($x = y$). Fractions of genome subsampled are indicated between the parentheses. Small population. **B:** Same as panel A but for the large population. **C:** Comparison of ROHs distributions from RAD-sequencing and the true distributions of IBD segments (defined as segments which coalesced less than 100 reproductive cycles ago). Black horizontal lines correspond to total IBD segments lengths per individuals (y axis) falling into the different length classes (x axis). Barplots show the mean (\pm sd) difference between the mean total length of IBD segment and their estimation for each sequencing method. Barplots below the horizontal black line indicate an underestimation while barplots above the horizontal black line indicate an overestimation of the total length of segments. Mean (and sd) are among individuals, simulation and subsampling replicates. Small population. **D:** Same as panel C but for the large population

Adapting the --homozyg-snps and --homozyg-window-snps parameters influences the estimation of F_{ROH} , leading even to a negative correlation between F_{ROH} and F_{IBD} for the large population and the smaller SNP density. These parameters do not seem to strongly influences the ROHs

MOLECULAR ECOLOGY RESOURCES

distributions estimation expect concerning the smaller ROHs class where we see a strong overestimation.

MOLECULAR ECOLOGY RESOURCES

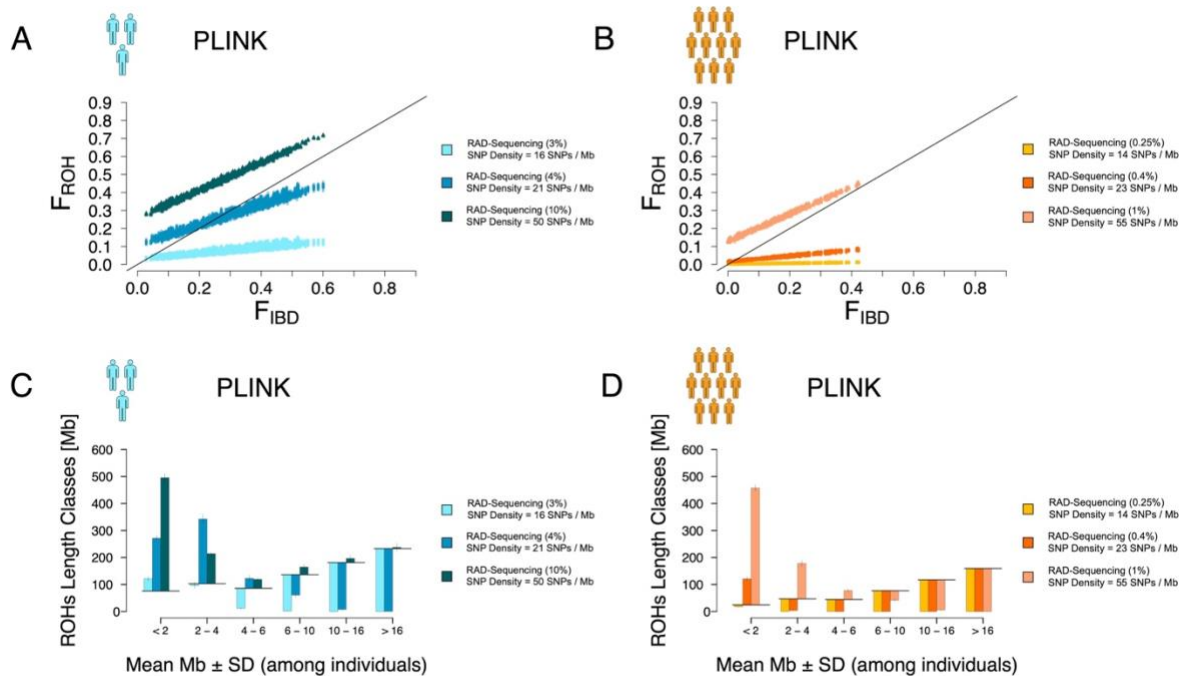


FIG S9: ROHs have been called with PLINK and a maximum length allowed between two adjacent SNPs (--homozyg-gap) of 200. --homozyg-window-snp and --homozyg-snp were dependent on the SNP density in the reduced dataset but with a minimum value of 30 and a maximum value of 100. **A:** Comparison between F_{ROH} estimated with RAD-sequencing on the y axis according to F_{IBD} (the true fraction of genome within IBD segments coalescing less than 100 reproductive cycles ago) on the x axis. Each point represents one individual (for one subsampling replicate within one simulation replicate). The black line represents the equality line ($x = y$). Fractions of genome subsampled are indicated between the parentheses. Small population. **B:** Same as panel A but for the large population. **C:** Comparison of ROHs distributions from RAD-sequencing and the true distributions of IBD segments (defined as segments which coalesced less than 100 reproductive cycles ago). Black horizontal lines correspond to total IBD segments lengths per individuals (y axis) falling into the different length classes (x axis). Barplots show the mean (\pm sd) difference between the mean total length of IBD segment and their estimation for each sequencing method. Barplots below the horizontal black line indicate an underestimation while barplots above the horizontal black line indicate an overestimation of the total length of segments. Mean (and sd) are among individuals, simulation and subsampling replicates. Small population. **D:** Same as panel C but for the large population

Reducing the --homozyg-gap parameter does not influence the estimation of F_{ROH} , but strongly influences the ROHs distributions estimation. No long ROHs are detected when this parameter is

MOLECULAR ECOLOGY RESOURCES

too small and there is a strong overestimation of small ROHs. Long IBD segments are probably split in small adjacent ROHs.

MOLECULAR ECOLOGY RESOURCES

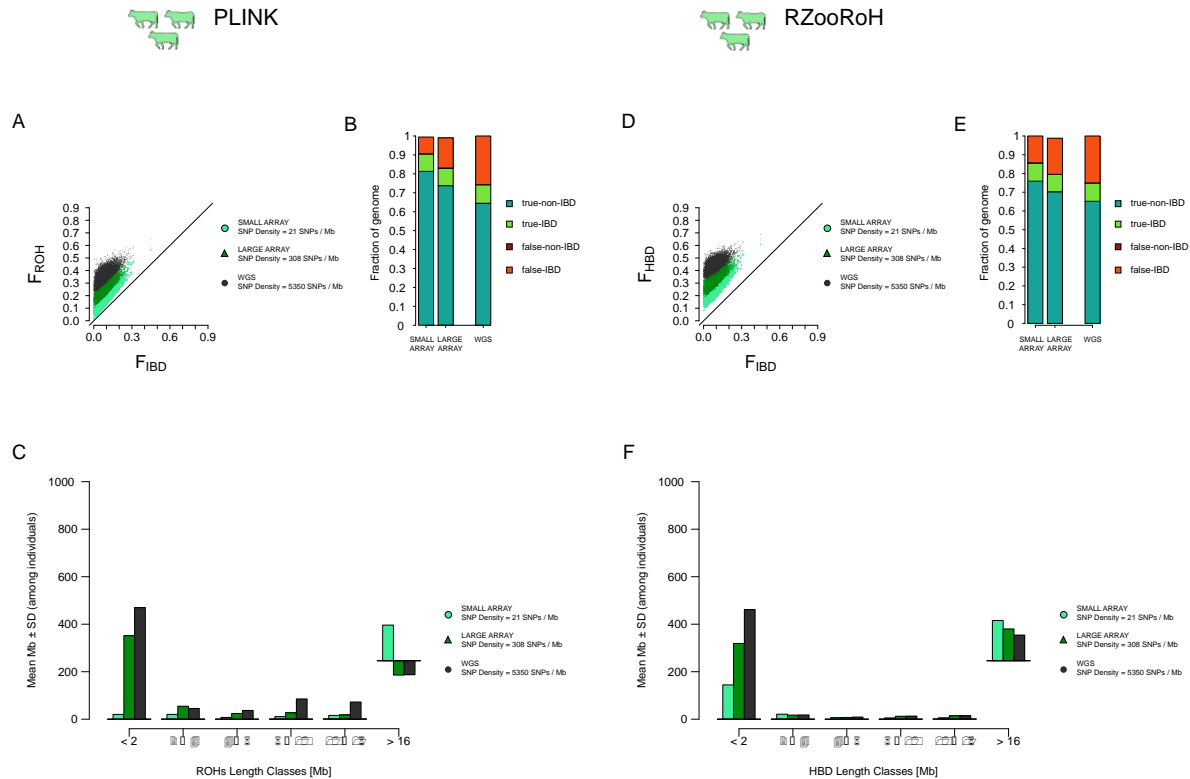

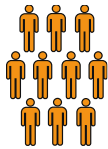


FIG S10: Comparison of ROHs and HBD segments detection with WGS and both SNPs arrays in the cattle population and for both PLINK and RZooRoH. **A:** Comparison of F_{ROH} estimated with PLINK with WGS and both SNPs arrays in regards to the true fraction of genome within IBD segments (coalescing less than 100 reproductive cycles ago). **B:** Fraction of genome correctly assigned outside IBD segments (true-non-IBD), correctly assigned within IBD segments (true-IBD), incorrectly assigned outside IBD segments (false-non-IBD) and incorrectly assigned within IBD segments (false-IBD) with WGS and both arrays. ROHs were called with PLINK **C:** Comparison of ROHs distributions with WGS and both arrays in regards to the true distribution of IBD segments (coalescing less than 100 reproductive cycles ago). Horizontal black lines represent the true mean (among simulation replicate) individual total IBD segments lengths estimated from simulated data for each length class. Barplots represent the mean (among simulation and subsampling replicate) difference between the estimated distributions (ROHs) and the truth (IBD segments). ROHs were called with PLINK. **D:** Comparison of F_{HBD} estimated with RZooRoH with WGS and both SNPs arrays in regards to the true fraction of genome within IBD segments (coalescing less than 100 reproductive cycles ago). **E:** Fraction of genome correctly assigned outside IBD segments (true-non-IBD), correctly assigned within IBD segments (true-IBD), incorrectly assigned outside IBD segments (false-non-IBD) and incorrectly assigned within IBD segments (false-IBD) with WGS and both arrays. HBD segments were called with RZooRoH **F:** Comparison of HBD segments distributions with WGS and both arrays in regards to the true distribution of IBD segments (coalescing less than 100 reproductive cycles ago). Horizontal black lines represent the true mean (among simulation replicate) individual total IBD segments lengths estimated from simulated data for each length class. Barplots represent the mean (among simulation and subsampling replicate) difference between the estimated distributions (HBD segments) and the truth (IBD segments). HBD segments were detected with RZooRoH.

MOLECULAR ECOLOGY RESOURCES

Table S3: Summary of the relationship between true fraction of genome within IBD segments (which coalesce less than 100 reproductive cycles ago) and inbreeding coefficients (F) estimated different sequencing methods: i) F_{ROH} estimated with PLINK ii) F_{HBD} estimated with RZooRoH and iii) F_{HOM} . For each replicate, the correlation is estimated as the Pearson correlation between F_{ROH} or F_{HBD} and F_{IBD} and the slope and intercept are extracted from the linear model: $F_{ROH \text{ OR } HBD} \sim F_{IBD}$. This table presents the mean (\pm sd) among replicates. For RAD-sequencing, the numbers in parentheses indicate the percentage of genome sequenced.

Sequencing method	F estimate	Correlation \pm sd	Intercept \pm sd	Slope \pm sd
RAD (0.05%)	F_{HBD}	0.932 \pm 0.013	-0.02 \pm 0.014	0.993 \pm 0.046
	F_{HOM}	0.897 \pm 0.014	-0.064 \pm 0.007	1.061 \pm 0.015
RAD (0.15%)	F_{HBD}	0.975 \pm 0.005	-0.01 \pm 0.009	0.987 \pm 0.031
	F_{HOM}	0.959 \pm 0.006	-0.065 \pm 0.007	1.062 \pm 0.016
RAD (2%)	F_{HBD}	0.995 \pm 0.001	0.045 \pm 0.005	0.945 \pm 0.014
	F_{HOM}	0.99 \pm 0.002	-0.065 \pm 0.007	1.062 \pm 0.016
Small array	F_{HBD}	0.996 \pm 0	0.193 \pm 0.004	0.804 \pm 0.014
Large array	F_{HBD}	0.997 \pm 0.001	0.203 \pm 0.004	0.79 \pm 0.014
 WGS	F_{HBD}	0.998 \pm 0	0.147 \pm 0.003	0.849 \pm 0.008
	RAD (3%)	F_{ROH}	-0.415 \pm 0.106	0.027 \pm 0.004
F_{HOM}		0.994 \pm 0.001	-0.065 \pm 0.007	1.063 \pm 0.016
RAD (4%)	F_{ROH}	0.984 \pm 0.005	0.129 \pm 0.009	0.672 \pm 0.051
	F_{HOM}	0.994 \pm 0.001	-0.065 \pm 0.007	1.062 \pm 0.016
RAD (10%)	F_{ROH}	0.996 \pm 0.001	0.262 \pm 0.005	0.766 \pm 0.01
	F_{HOM}	0.995 \pm 0.001	-0.065 \pm 0.007	1.063 \pm 0.016
Small array	F_{ROH}	0.997 \pm 0	0.091 \pm 0.003	0.891 \pm 0.012
Large array	F_{ROH}	0.998 \pm 0	0.135 \pm 0.003	0.86 \pm 0.012
WGS	F_{ROH}	0.999 \pm 0	0.087 \pm 0.003	0.918 \pm 0.006
RAD (0.002%)	F_{HBD}	0.846 \pm 0.035	0.106 \pm 0.018	0.891 \pm 0.071
	F_{HOM}	0.758 \pm 0.022	-0.014 \pm 0.002	1.017 \pm 0.011
 RAD (0.008%)	F_{HBD}	0.914 \pm 0.019	0.104 \pm 0.012	0.892 \pm 0.05
	F_{HOM}	0.851 \pm 0.017	-0.013 \pm 0.001	1.009 \pm 0.007
RAD (0.125%)	F_{HBD}	0.993 \pm 0.002	0.104 \pm 0.003	0.891 \pm 0.012
	F_{HOM}	0.987 \pm 0.002	-0.013 \pm 0.001	1.011 \pm 0.006
Small array	F_{HBD}	0.999 \pm 0	0.047 \pm 0.001	0.946 \pm 0.006
Large array	F_{HBD}	0.999 \pm 0	0.212 \pm 0.001	0.782 \pm 0.004

MOLECULAR ECOLOGY RESOURCES

WGS	F_{HBD}	0.999 ± 0	0.171 ± 0.001	0.825 ± 0.004
RAD (0.25%)	F_{ROH}	0.241 ± 0.149	0.018 ± 0.002	0.009 ± 0.006
	F_{HOM}	0.993 ± 0.001	-0.013 ± 0.001	1.012 ± 0.006
RAD (0.4%)	F_{ROH}	0.993 ± 0.002	0.101 ± 0.005	0.831 ± 0.02
	F_{HOM}	0.995 ± 0.001	-0.013 ± 0.001	1.012 ± 0.005
RAD (1%)	F_{ROH}	0.998 ± 0.001	0.168 ± 0.005	0.864 ± 0.008
	F_{HOM}	0.998 ± 0	-0.013 ± 0.001	1.012 ± 0.005
Small array	F_{ROH}	0.999 ± 0	0.014 ± 0.001	0.96 ± 0.005
Large array	F_{ROH}	0.999 ± 0	0.065 ± 0.001	0.932 ± 0.005
WGS	F_{ROH}	0.999 ± 0	0.036 ± 0.001	0.963 ± 0.005

MOLECULAR ECOLOGY RESOURCES

REFERENCES

- Ceballos, F. C., Hazelhurst, S., & Ramsay, M. (2018). Assessing runs of Homozygosity: A comparison of SNP Array and whole genome sequence low coverage data. *BMC Genomics*, *19*(1), 106. <https://doi.org/10.1186/s12864-018-4489-0>
- Chadeau-Hyam, M., Hoggart, C. J., O'Reilly, P. F., Whittaker, J. C., De Iorio, M., & Balding, D. J. (2008). Fregene: Simulation of realistic sequence-level data in populations and ascertained samples. *BMC Bioinformatics*, *9*(1), 364. <https://doi.org/10.1186/1471-2105-9-364>
- Duntsch, L., Whibley, A., Brekke, P., Ewen, J. G., & Santure, A. W. (2021). Genomic data of different resolutions reveal consistent inbreeding estimates but contrasting homozygosity landscapes for the threatened Aotearoa New Zealand hihi. *Molecular Ecology*, *30*(23), 6006–6020. <https://doi.org/10.1111/mec.16068>
- Ferenčaković, M., Hamzić, E., Gredler, B., Solberg, T. R., Klemetsdal, G., Curik, I., & Sölkner, J. (2013). Estimates of autozygosity derived from runs of homozygosity: Empirical evidence from selected cattle populations. *Journal of Animal Breeding and Genetics*, *130*(4), 286–293. <https://doi.org/10.1111/jbg.12012>
- Frantz, L. A. F., Bradley, D. G., Larson, G., & Orlando, L. (2020). Animal domestication in the era of ancient genomics. *Nature Reviews Genetics*, *21*(8), 449–460. <https://doi.org/10.1038/s41576-020-0225-0>
- Haller, B. C., & Messer, P. W. (2019). SLiM 3: Forward Genetic Simulations Beyond the Wright–Fisher Model. *Molecular Biology and Evolution*, *36*(3), 632–637. <https://doi.org/10.1093/molbev/msy228>
- Kelleher, J., Etheridge, A. M., & McVean, G. (2016). Efficient Coalescent Simulation and Genealogical Analysis for Large Sample Sizes. *PLOS Computational Biology*, *12*(5), e1004842. <https://doi.org/10.1371/journal.pcbi.1004842>
- Meyermans, R., Gorssen, W., Buys, N., & Janssens, S. (2020). How to study runs of homozygosity using PLINK? A guide for analyzing medium density SNP data in livestock and pet species. *BMC Genomics*, *21*(1), 94. <https://doi.org/10.1186/s12864-020-6463-x>
- Purfield, D. C., Berry, D. P., McParland, S., & Bradley, D. G. (2012). Runs of homozygosity and population history in cattle. *BMC Genetics*, *13*(1), 70. <https://doi.org/10.1186/1471-2156-13-70>
- Purfield, D. C., McParland, S., Wall, E., & Berry, D. P. (2017). The distribution of runs of homozygosity and selection signatures in six commercial meat sheep breeds. *PLOS ONE*, *12*(5), e0176780. <https://doi.org/10.1371/journal.pone.0176780>
- Qanbari, S., & Wittenburg, D. (2020). Male recombination map of the autosomal genome in German Holstein. *Genetics Selection Evolution*, *52*(1), 73. <https://doi.org/10.1186/s12711-020-00593-z>
- Stoffel, M. A., Johnston, S. E., Pilkington, J. G., & Pemberton, J. M. (2021). Genetic Architecture and Lifetime Dynamics of Inbreeding Depression in a Wild Mammal. *Nature Communications*, *12*(1), 2972. <https://doi.org/10.1038/s41467-021-23222-9>

Supplementary Material Chapter II: Detecting inbreeding depression in structured populations

Authors: Eléonore Lavanchy, Bruce S. Weir and Jérôme Goudet

Detailed introduction

Inbreeding depression, the decrease in mean phenotypic value in inbred individuals, is a phenomenon pervasive in humans, domestic and wild animals and plants [55, 25, 14]. Inbreeding has been associated with various diseases and is the result of mating between relatives. It has been observed in a wide range of taxa such as humans [2, 10], livestock [41, 28], wild animal populations [25] and plants [25, 57]. The effect of inbreeding on individuals' genomes is to increase homozygosity: since related individuals share higher genetic similarity, their offspring are likely to harbor higher fractions of homozygous identical-by-descent (IBD) genomic regions. On the one hand, mating between closely related individuals such as siblings or first-degree cousins results in strong inbreeding, also referred to as 'recent inbreeding'. For instance, the mating of closely related individuals is encouraged in domestic species, as part of the artificial selection of the process of domestication is mating individuals with similar phenotypes of economic interest [35, 40]. Nonetheless, recent inbreeding also occurs in wild isolated populations especially those with extremely small effective sizes [24, 13] and in various human populations where mating with relatives is culturally encouraged [3]. On the other hand, more moderate and remote inbreeding may occur because of ancient relatedness between the parents, which is often observed in populations with small effective sizes or ancient founder effect. This type of inbreeding has been observed in humans, notably in the individuals with East Asian ancestry from the 1,000 Genomes Project [17] but also happens in wild [47] and domestic [35] populations with small sizes.

What is inbreeding depression ?

Inbreeding is often associated with reduced fitness, a phenomenon called inbreeding depression (ID), in many different species such as humans [9, 8], other animals [44, 20, 33], and plants [57]. Charlesworth and Willis suggested two specific mechanisms by which increased homozygosity leads to a reduction in population fitness [12]. The first mechanism is heterozygous advantage where heterozygotes are at an advantage. Inbred

individuals have higher chances of being homozygous, thus they will tend to have lower fitness. This mechanism has been characterized in *Drosophila* [12] but is likely to be of lesser importance in other species [12, 19]. The second mechanism involves partially recessive deleterious alleles which, when only present in one copy, have little effect and are therefore undetected by natural selection. Hence, selection can only act upon them when they are in the homozygous state. For partially recessive deleterious alleles, the strength of selection is weak, even in the homozygous state. As a result, many of these loci are segregating at low frequencies in populations [43]. For non-inbred individuals, the effect on fitness is minimal: they might carry some of these alleles in homozygous form by chance, but since most of these alleles are only partly detrimental, the individuals' overall fitness is hardly impacted. Conversely, the more inbred an individual is, the higher is its proportion of genome in the homozygous state, increasing the likelihood of a large quantity of partially deleterious alleles in a homozygous state. As these alleles accumulate, their deleterious effects will have a significant impact on the fitness of inbred individuals. In populations with small sizes, these marginally recessive alleles can easily reach intermediate frequencies [25] or even fixation [16] since the effect of drift will be much stronger. Moreover, individuals will tend to share more co-ancestry in these small populations, resulting in a large accumulation of these deleterious recessive alleles in the homozygous state.

Why measuring inbreeding ?

Since it can have disastrous effects on populations, quantifying inbreeding and its deleterious consequences is of the utmost importance. In humans, for example, researchers were able to link inbreeding with many deleterious phenotypes [51, 42] which led to a better understanding of the underlying mechanisms involved in these traits. Quantifying inbreeding is also essential for monitoring endangered and small isolated populations. If their inbreeding status is high, we can expect a decline of the population in future generations. To avoid this issue, strict breeding programs which aim at reducing the overall inbreeding load of the population can be implemented [29, 46, 48].

How to quantify inbreeding ?

Many different methods have been developed for inbreeding quantification and there is no consensus on which one is the best [1, 7, 18, 39, 56, 58]. The classical approach was first proposed by Sewall Wright in 1922 and makes use of pedigrees (called hereafter F_{PED})

[53]. F_{PED} corresponds to the probability that two alleles are IDB (a definition proposed by Malécot in 1948 [32]) and rely on the genealogy of the population. Consequently, its estimation is only possible in populations where matings are actively recorded (i.e. mainly human and domestic populations). Furthermore, what F_{PED} measures is the expected inbreeding coefficient, which can be very different from the realized coefficient due to recombination stochasticity and random segregation of alleles. With the advances in sequencing technologies, genomic-based inbreeding coefficients (hereafter called $F_{genomic}$) have been developed. Among these, some coefficients rely on the comparison between observed and expected heterozygosity such as F_{HOM} [11, 45], the expected allele sharing between individuals such as F_{AS} [58] or on the correlation between uniting gametes such as F_{UNI} [54]. In addition to estimating the realized inbreeding coefficient and requiring no prior knowledge of the mating behavior of the population, these genomic estimates are simple and straightforward to compute and do not require whole-genome sequencing (WGS) data; a few thousands SNPs are usually sufficient for reliable inbreeding estimation in humans [18]. However they also have a disadvantage: they usually rely on allelic frequencies (except for F_{AS}) and therefore if these frequencies have not been correctly estimated, this will affect the estimation of these coefficients. Additionally, these coefficients treat each SNP independently, whereas in nature DNA is transmitted from parents to offspring in large chromosomal chunks. In order to take this into account, McQuillan *et al.* (2008) proposed a new inbreeding coefficient: F_{ROH} which uses runs of homozygosity (ROHs) long homozygous stretches as proxy for IBD segments within individuals [36]. A model-based approach relying on hidden Markov models (HMM) has also been developed for detecting IBD segments [31] by identifying homozygous-by-descent (HBD) segments. This model is the basis for many other model-based IBD segments detection methods such as `BCFTTOOLS` [38], `BEAGLES` [4] and `RZOOROH` [15]. The inbreeding coefficient estimated with these model-based approaches will be called F_{HBD} from now on. One advantage of these methods is that they do not rely on allelic frequencies which can be very valuable when only a few individuals are available. However, it has been shown that these coefficients and especially F_{ROH} are sensitive to SNPs density and parameters sets and no consensus on what is the best set of parameters exists nowadays [37, 30].

How to quantify inbreeding depression

How to quantify inbreeding depression, although central to conservation genetics for decades [25] is still debated. This debate includes two sub-questions: which statistical model should be utilized ? And which inbreeding coefficient ? Regarding the model, the classical approach consisted of using linear regression of the phenotypes on the inbreeding coefficient. However, other models have been used, such as maximum likelihood and Generalized Linear models (GLMs) with various link functions. In 2019, Nietlisbach *et al.* [39] compared different models and found that the common GLM models with logit link did not allow for accurate inbreeding depression strength estimation. They propose using maximum likelihood estimation or GLM with logarithm link functions.

Effect of sample size

Except for humans and domestic species where genetic and phenotypic data are available for several thousands of individuals [56, 50, 34, 44, 49], ID studies in the wild are usually performed on smaller sample sizes varying between 100 and few thousands individuals [20, 25, 26, 23]. Consequently, we may not be able to detect inbreeding depression in many wild populations (unless the effect is very strong). Indeed Keller *et al.* [27] stressed that to detect the effect of deleterious alleles with small effects (and F_{ROH}), very large sample sizes of thousands of individuals are needed.

Review of what has been done so far

An unresolved issue is which inbreeding coefficient is more accurate for quantifying inbreeding depression. In 2011, Keller *et al.* [27] performed simulations mimicking past human demography and compared different inbreeding coefficients. The authors showed that F_{ROH} retains more individual variation compared to SNPs-independent measures of inbreeding and correlates best with homozygous mutation load which they suggest is likely to make it the best F for quantifying inbreeding depression. However, the authors stress that to detect the effect of deleterious alleles with small effects, very large sample sizes are needed. In 2015, Kardos *et al.* [21] also performed simulations and compared the ability of F_{PED} , F_{ROH} and F_{HOM} to capture the true proportion of genome within IBD segments. They found similar results to [27]: F_{PED} is outperformed by all $F_{genomic}$ and among the $F_{genomic}$ they tested, F_{ROH} performed better. In 2016, Bérénos *et al.* [5] used pedigrees and several $F_{genomic}$ and showed that genomic based coefficients of inbreeding detect more

inbreeding depression compared to F_{PED} . In 2017, Yengo *et al.* [56] used a homogeneous subset of the UK biobank dataset (individuals of European ancestries exclusively, and with kinships less than 0.05) to simulate traits and compare various F . The results they found contradicted Kardos *et al.* [21]: they found F_{UNI} to be the best coefficient to estimate ID and that F_{ROH} -based estimates of ID tended to be overestimated and showed higher standard error. In addition, one aspect discussed in this paper but not elsewhere as far as we know, is the spurious effect that directional additive effect can have on estimates of inbreeding depression (what the authors called DEMA, for Directional Effect of Minor Alleles). For a trait linked to fitness, we expect most new mutations to have detrimental effects, diminishing the value of the trait. Selection will tend to remove these detrimental alleles or maintain them at low frequencies. Many low frequency alleles would then be detrimental, leading to a negative DEMA. The authors showed that F_{HOM} (and thus F_{AS} since they have similar properties) is sensitive to DEMA while F_{UNI} and F_{ROH} are not. They also showed via simulations that all estimates of inbreeding depression are somewhat sensitive to population structure, F_{UNI} being the least affected. They recommend estimating inbreeding using Linkage Disequilibrium (LD) score and Minor Allele Frequency (MAF) bins, and to sum the ID estimates from these bins as an overall estimate of ID for the trait. In response to this article, Kardos and coauthors [22] argued that F_{UNI} yielded better results than F_{ROH} because of the method Yengo *et al.* [56] used to compare the performance of the different F s and that F_{ROH} is preferable for studying inbreeding depression. In 2019, Nietlisbach *et al.* [39] published a paper using simulations and compared the capacity of different F s to quantify ID. They used the inbreeding load as the gold standard and found that F_{ROH} was the coefficient which showed the highest correlation with inbreeding load. In 2020, Caballero *et al.* [7] used simulations and included several populations with different histories: they found that the best F actually depends on the size of the population. F_{ROH} did a better job at quantifying ID in population with small effective size while F_{UNI} was better at predicting ID estimates in populations with large effective sizes. However, the authors stressed that the error in ID estimates were large in all situations. Finally, in 2021, Alemu *et al.* [1] used SNPs-array empirical cattle data for several groups of allelic frequencies and found that F_{UNI} and F_{GRM} are better at quantifying homozygosity at rare alleles while F_{ROH} and F_{HOM} are better for alleles at intermediate frequencies and correlate better with whole-genome homozygosity. Consequently, the authors suggest that the history of the population will play a key role in determining which F is best to estimate ID. Indeed, recessive deleterious alleles which should be those responsible for inbreeding depression are expected to segregate at low

frequencies in large populations due to negative selection. On the contrary, in small populations, drift can increase these deleterious recessive alleles frequencies to reach intermediate frequencies which would make F_{ROH} and F_{HOM} better suited to detect ID.

Summary of what we did, found and propose

In this paper we simulated traits based on simulated as well as empirical WGS human data from populations with various sizes from the 1,000 Genomes project. We show that some F are more sensitive to population structure and DEMA than others. We confirm only some of Yengo *et al.* [56] results. Importantly, we show that accounting for the non-independence of observations with a mixed model via an allele-sharing based genomic relationship matrix (GRM) and using a modified version of F_{UNI} which gives more weight to common alleles resolves most of the issues raised by Yengo *et al.* [56].

Supplementary Material and methods

Summary of the simulated scenarios

Table S1 presents the different scenarios shown in these SM and the parameters associated with each scenario: namely whether the additive and dominance effect sizes were randomly assigned or proportional to MAF as well as whether the simulation scenario included DEMA (Directional Effect of Minor Alleles). Figure S1 depicts a graphical representation of the three parameters mentioned above.

Table S1: Simulated scenarios used in this study. The first column corresponds to the scenario name and the three others indicate whether additive effect sizes (second column) and dominance coefficients (third column) were randomly assigned or according to MAF and whether DEMA (third column) was included.

Scenario	Additive effect sizes	Domiannce effect sizes	DEMA
Standard	Randomly assigned	Randomly assigned	No
ADD	Proportional to MAF	Randomly assigned	No
DOM	Randomly assigned	Proportional to MAF	No
DEMA	Randomly assigned	Randomly assigned	Yes
ADD & DOM	Proportional to MAF	Proportional to MAF	No
ADD & DEMA	Proportional to MAF	Randomly assigned	Yes
DOM & DEMA	Randomly assigned	Proportional to MAF	Yes
ADD & DOM & DEMA	Proportional to MAF	Proportional to MAF	Yes
ADD & DOM & DEMA	Proportional to MAF	Proportional to MAF	Yes

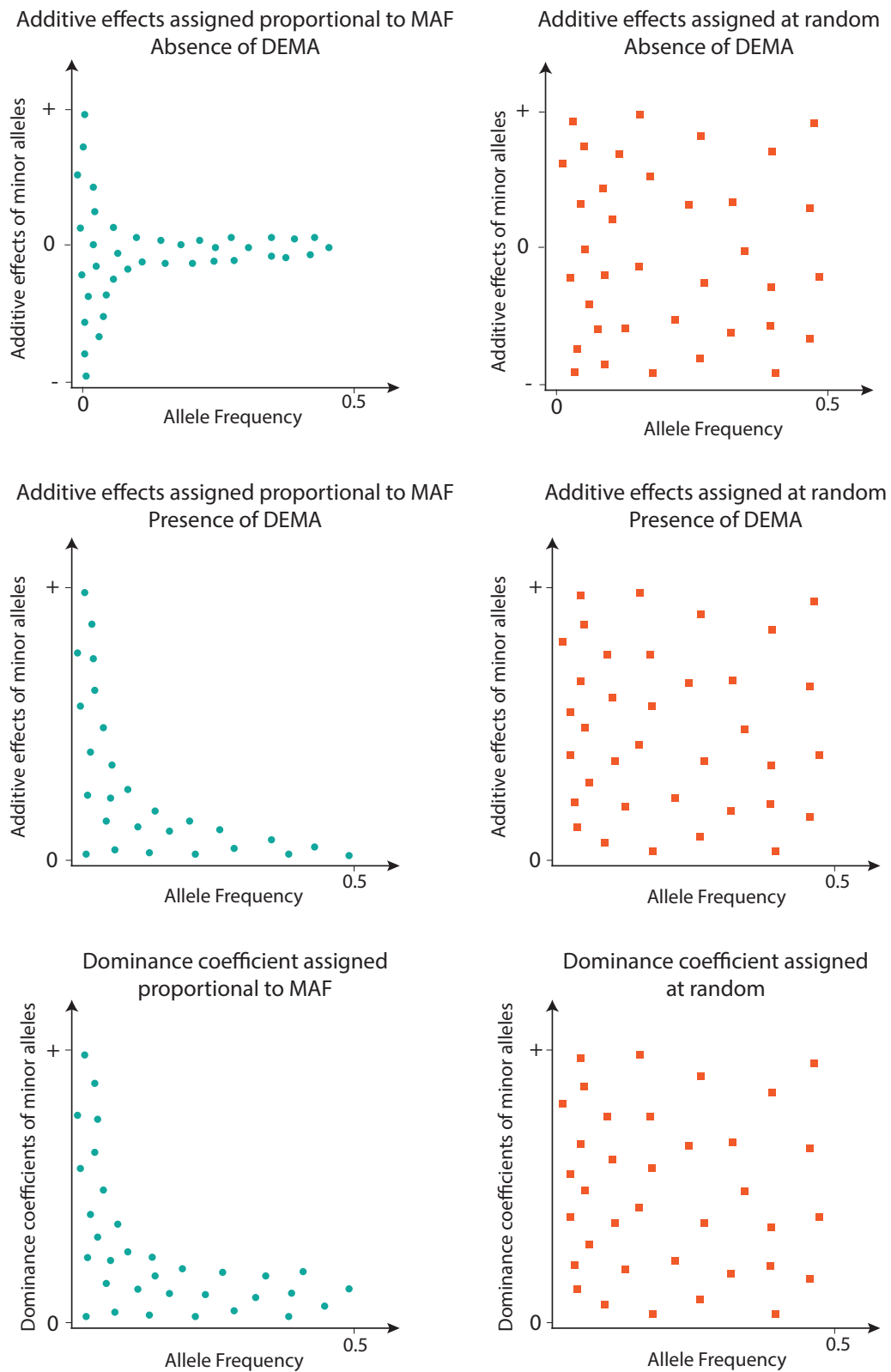


Figure S1: Distribution of the additive effect sizes and dominance and dominance coefficients when proportional to MAF or randomly assigned in the absence and presence of DEMA.

LDMS Stratification

For the SNPs-based inbreeding coefficients (i.e. F_{AS} and both F_{UNI}), we also investigated whether using the Linkage disequilibrium and minor allele frequency stratified inference (hereafter called LDMS stratification) proposed by Yengo *et al.* (2017) [56] improved the estimation of the inbreeding depression strength (b). For LDMS stratification, each F is estimated from each combination of 7 MAF bins and 4 LD bins, and the ID estimate is obtained as the sum of the partial regression coefficients of the trait on each of the 28 inbreeding coefficients. The MAF and LD bins are defined as in [56]: MAF₁: $p \leq 0.001$; MAF₂: $0.001 < p \leq 0.01$; MAF₃: $0.01 < p \leq 0.1$; MAF₄: $0.1 < p \leq 0.2$; MAF₅: $0.2 < p \leq 0.3$; MAF₆: $0.3 < p \leq 0.4$; MAF₇: $0.4 < p \leq 0.5$, and LD bins correspond to the 4 quartiles. For EAS samples, since there are only 500 samples, we discarded MAF₁ and made MAF₂: $p \leq 0.01$. Results for these analyses can be found in figures S10 – S17.

MAF filtering

In order to verify whether rare alleles were responsible for F_{UNI}^u poor estimations of b , we did as Yengo *et al.* (2017) [56] and filtered the WORLD populations genomic data to keep only SNPs with $MAF > 0.05$ using `BCFT00LS`. We then re-estimated all F and GRMs on the newly filtered data set and re-simulated inbreeding depression. Results for this analysis can be found in figure S18.

HBD segments without size selection

When estimating IBD segments in the genome, an advantage of model-based approaches (such as `BCFT00LS`) is that there is no constraint on the minimum size of an HBD segment. However, since the coalescing events responsible for inbreeding depression are usually recent, we only considered segments larger than 100KB and 1MB in the main text. In supplementary figure S19, we also try to estimate ID with ROHs and HBD segments larger than 5Mb. We wanted to test whether using all HBD segments independently of their size would yield better b estimates. Consequently, in the WORLD population and from the output of `BCFT00LS`, we did not filter on size but rather on quality score. Indeed, the `BCFT00LS` output includes a quality score which gives an indication about how confident we are about an HBD segment being IBD. A minimum quality of 30 was used for filtering. Results for this analysis can be found in figure S19.

House sparrow data set analysis

A metapopulation of house sparrows (*Passer domesticus*) from several islands in Northern Norway has been monitored since 1993 and Niskanen *et al.* [60] investigated inbreeding depression on several traits and made available phenotype and genotype data on more than 3,100 adult individuals. The data set is ideal to illustrate our method as individuals belong to many islands and the data is slightly genetically structured (global $F_{ST} = 0.028$) and some individuals are highly related (14,288 pairs ($\approx 1\%$) have allele-sharing kinships larger than 0.1, among which 3,453 ($\approx 0.2\%$) are larger than 0.2).

We used only morphological phenotypes, namely adult tarsus length, wing length, bill depth, bill length and mass, as they can be analysed with linear models. We removed information from non-autosome (scaffold 32) but otherwise kept all SNPs to avoid biases when filtering for minor allele frequencies and LD [59]. We filtered out individuals who were not present as adults in one of the eight studied islands, as was done in the original analysis [60]. The data set used for analysis contained 1,786 individuals genotyped at 181,529 SNPs. We compared the results of a simple linear model with Sex and F_{UNI}^W as explanatory variables, to the LMM_{AS} linear mixed model with Sex and F_{UNI}^W as fixed effects. We also analysed two additional linear mixed models to match the analyses carried out in Niskanen *et al.* [60], one with islands and years nested in islands as random effects, and a mixed model with these two random effects as well as the allele-sharing GRM. Niskanen *et al.* [60] ran a model with the pedigree relatedness matrix as random factor, but not with a GRM. Linear models were carried out with the *lm* function of R, while the mixed models were carried out with the *lmer* function of the *lme4* package or the *lmm.aireml* function of the *gaston* package if the model contained a GRM. The *lmm.aireml* function accepts random factors only if they are in matrix form, we thus converted the factors *islands* and *years:islands* to matrices made of 0s and 1s, 0 if the two observations did not belong to the same group and 1 if they did. We checked that *lmer* and *lmm.aireml* gave the same results in models with no GRMs. Finally, to test if b , the slope associated with F_{UNI}^W , was significantly different from 0, we used the *score.fixed.linear* function of the *gaston* package for models with GRMs.

Supplementary Results and Discussion

The following section contains all the supplementary figures and tables mentioned in the main text.

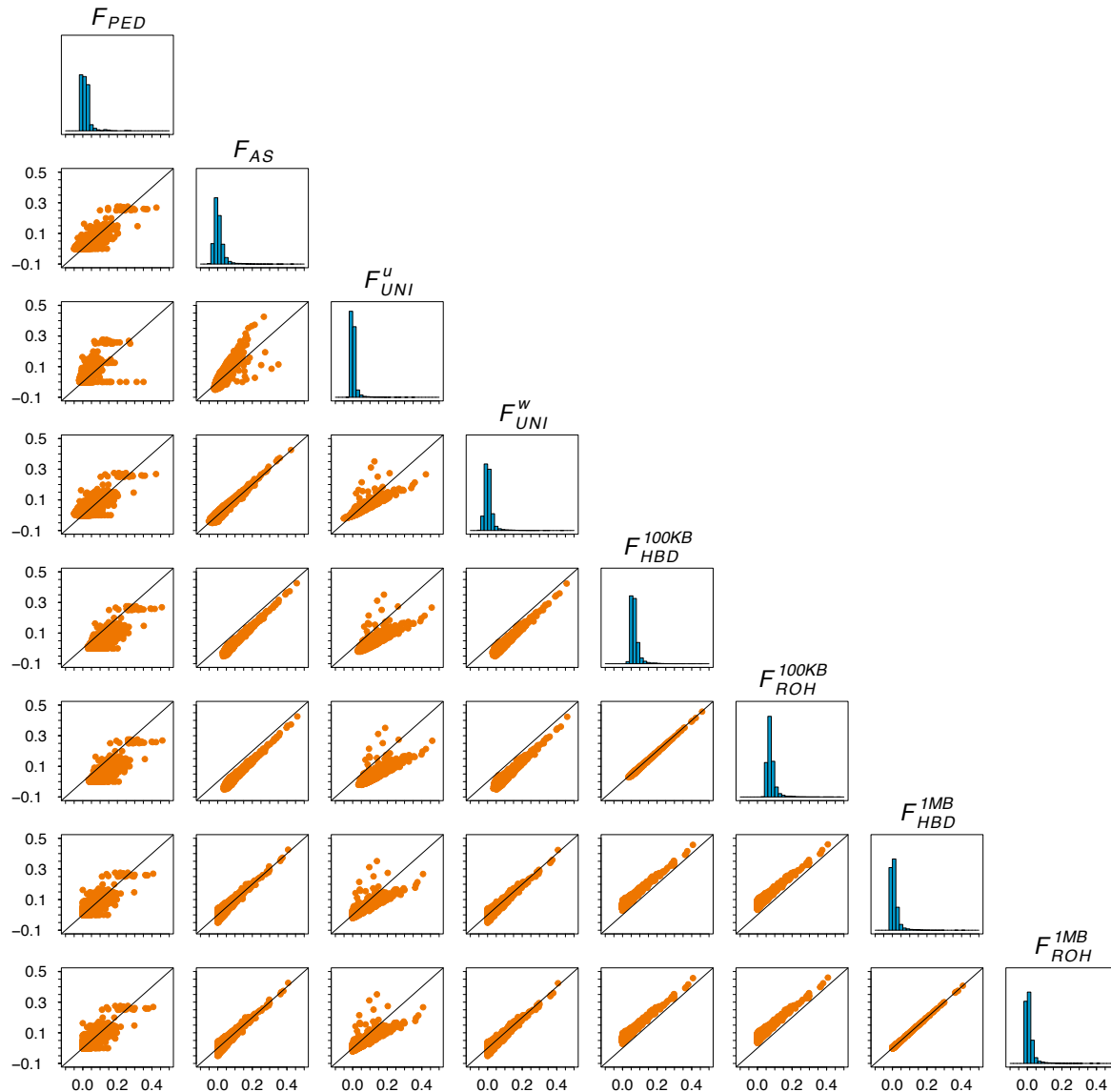


Figure S2: Pairwise comparison among the different inbreeding estimates (F) in the simulated PEDIGREE population. F depicted in this figure are F_{PED} , F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$ and $F_{ROH.1MB}$.

Figure S2 shows the comparison among the different inbreeding coefficients used in the PEDIGREE population. In general, we can see that there is a strong correlation between all $F_{genomic}$, with the exception of F_{UNI}^u and the other F . In addition, the non-genomic-based F_{PED} has the lowest correlation with all the other F .

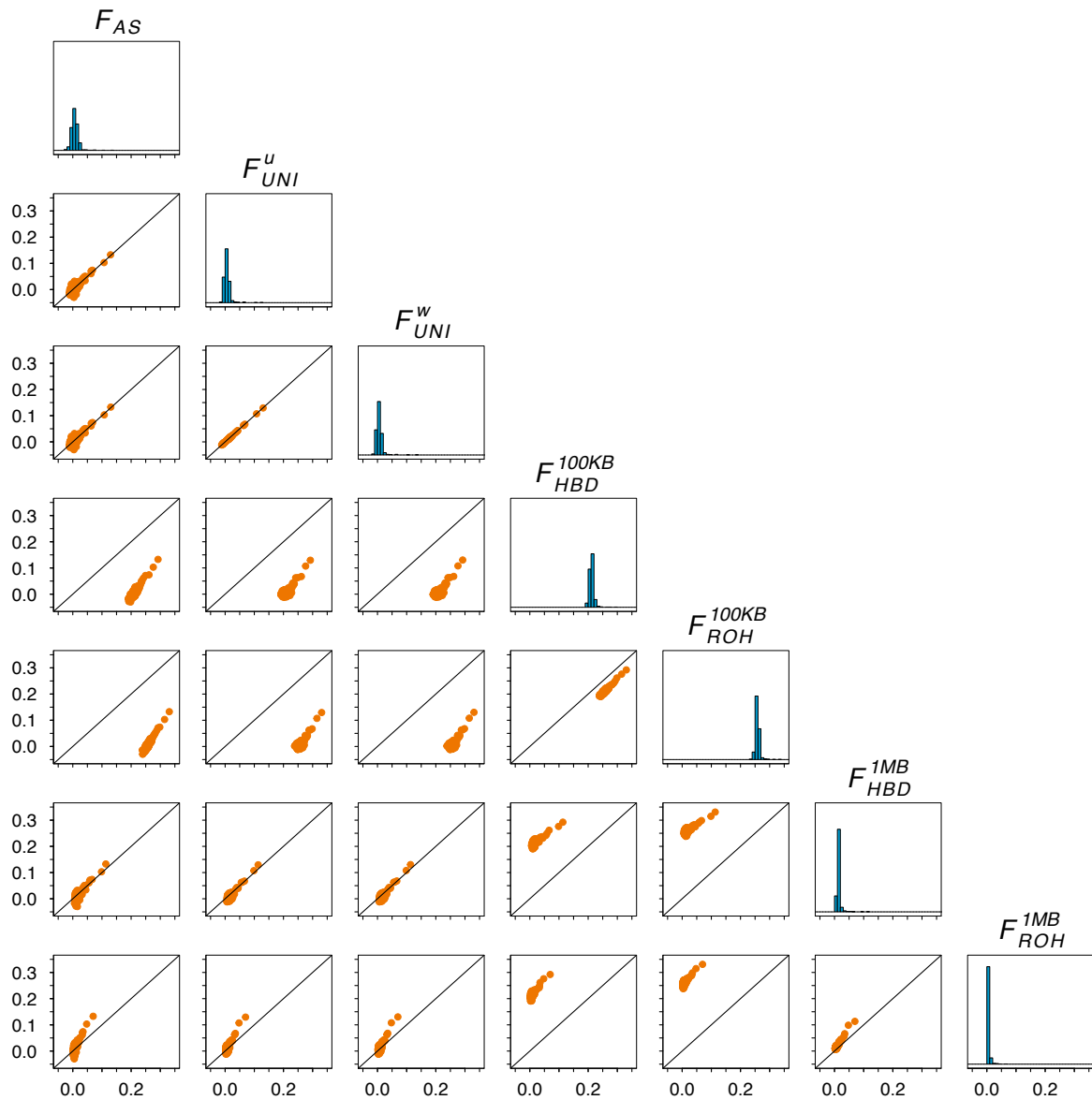


Figure S3: Pairwise comparison among the different inbreeding estimates (F) in the 1,000 Genomes Project EAS population. F depicted in this figure are F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$ and $F_{ROH.1MB}$.

Figure S3 shows the comparison among the different inbreeding coefficients used in the EAS population. With no structure in the population, there is a good correlation between all $F_{genomic}$. Even though the absolute values of the F are different, the rank of inbreeding is always conserved among individuals (i.e. the most inbred individuals are the most inbred for all F).

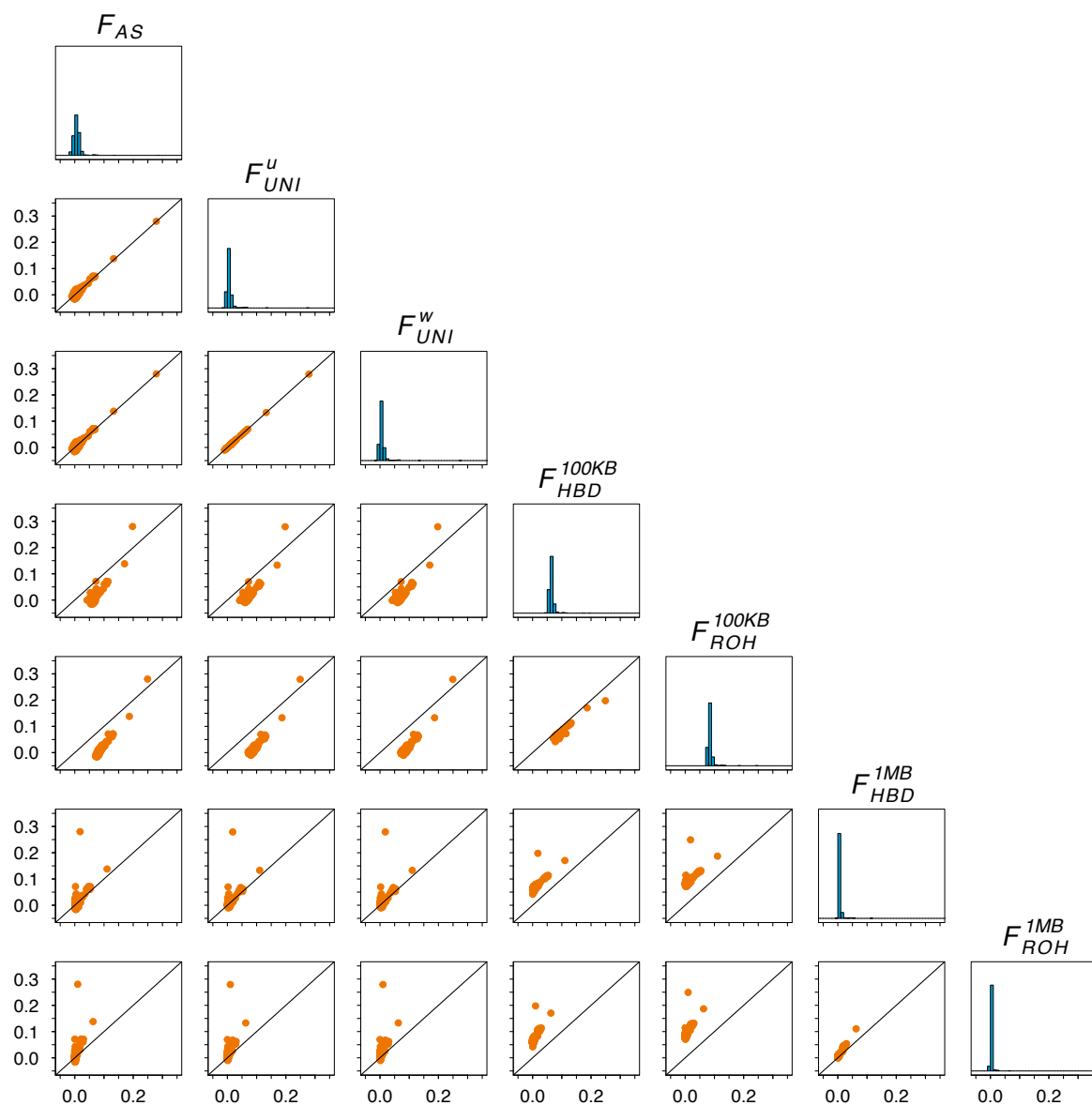


Figure S4: Pairwise comparison among the different inbreeding estimates (F) in the 1,000 Genomes Project AFR population. F depicted in this figure are F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$ and $F_{ROH.1MB}$.

Figure S4 shows the comparison among the different inbreeding coefficients used in the AFR population. In this population too, there is a good correlation between all $F_{genomic}$ (especially the three SNPs-based F : F_{AS} , F_{UNI}^u and F_{UNI}^w).

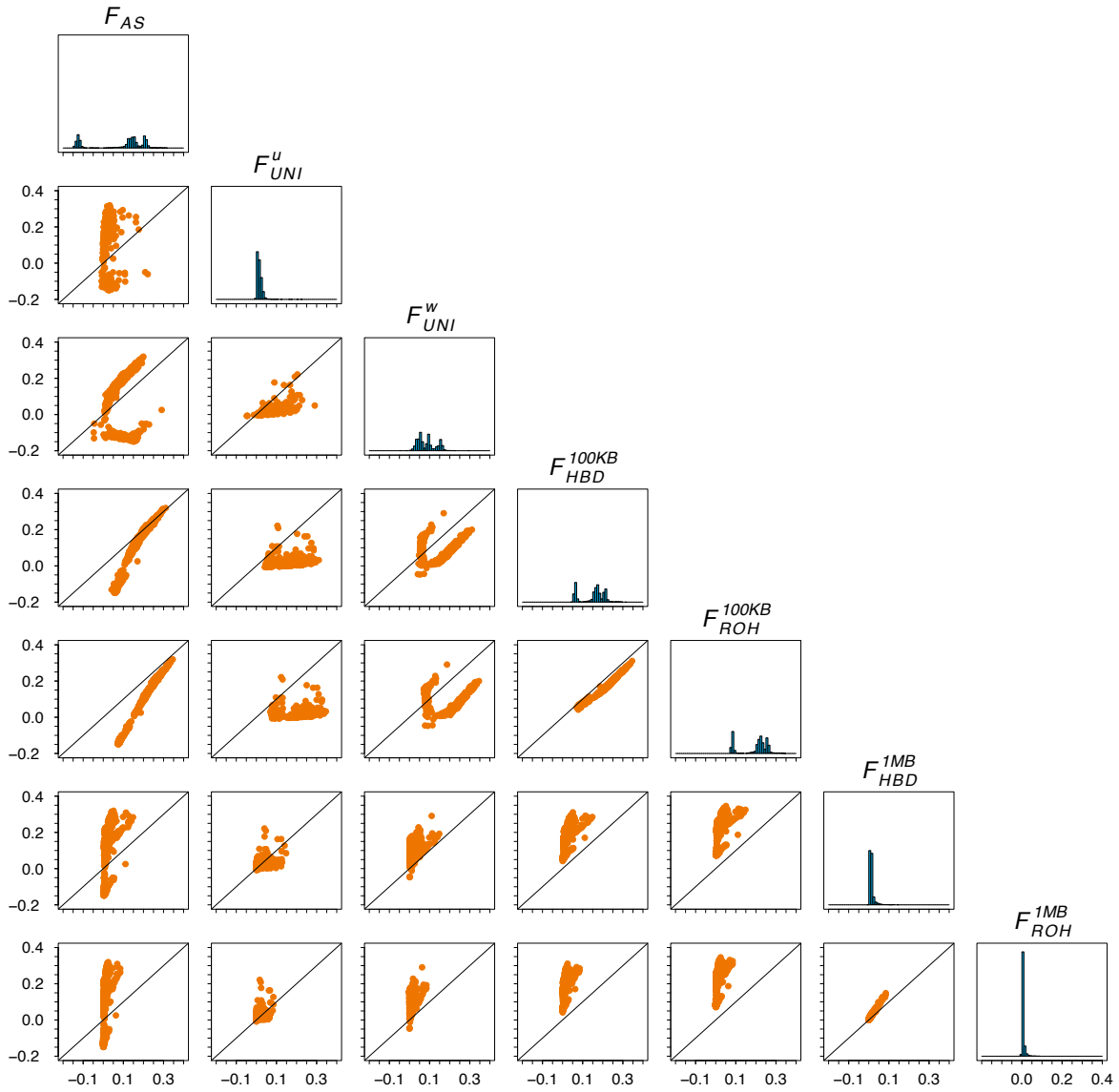


Figure S5: Pairwise comparison among the different inbreeding estimates (F) in the 1,000 Genomes Project WORLD population (all the individuals). F depicted in this figure are: F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$ and $F_{ROH.1MB}$.

Figure S5 shows the comparison among the different inbreeding coefficients used in the structured WORLD population. In this population as well, there is a good correlation between all F_{genomic} except F_{UNI}^u and the others. In addition, among the other F , the African samples are the ones with the lowest correlation.

LMM^w_{GCTA}

LMM^u_{GCTA}

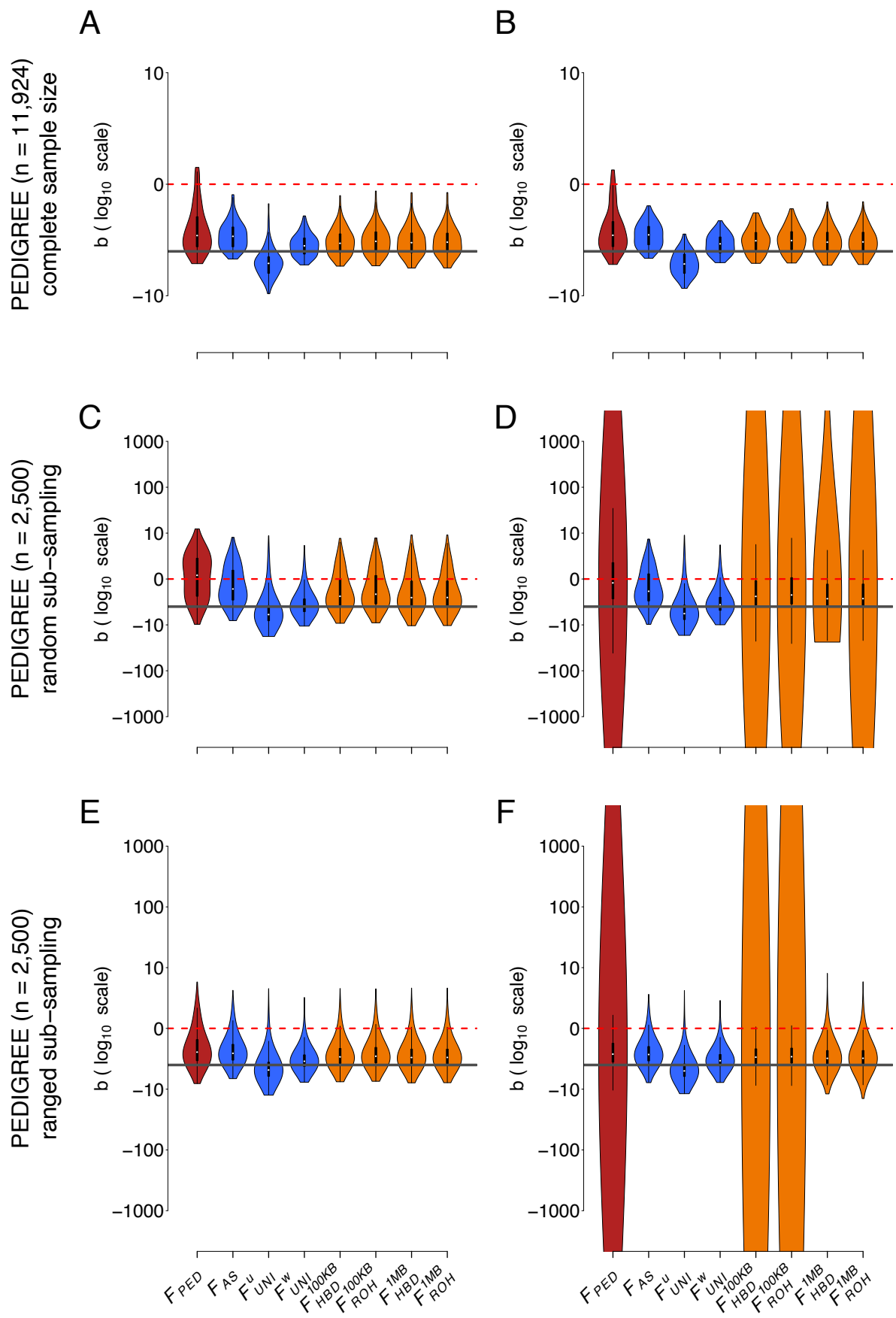


Figure S6: Comparison of the estimation of inbreeding depression strength (b) among different F estimates and the LMM including two different GRMs in the PEDIGREE population and with the ADD & DOM & DEMA scenario. The first column depicts the LMM including the $GCTA^w$ matrix (panels A, C and E) and the second column the linear mixed model with the $GCTA^u$ matrix (panels B, D and F). The first row shows the complete simulated population ($n = 11,924$ individuals) on panels A and B. The second row depicts the randomly subsampled population ($n = 2,500$ individuals) on panels C and D and the third row shows the ranged subsampled PEDIGREE population ($n = 2,500$ individuals) on panels E and F. Inbreeding estimates presented in this graph are F_{PED} , F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$ and $F_{ROH.1MB}$. For panels A and B, violin plots represent the distribution of the inbreeding depression strength estimates (b) among the simulated 100 replicates. For panels C to F, violin plots represent the distribution of the inbreeding depression strength estimates (b) among the 10,000 simulated and subsampling replicates (100 subsampling replicate for each of the 100 simulation replicates). The solid dark grey line is the true strength of ID ($b = -3$). The dashed red line represents the absence of ID ($b = 0$), meaning that we failed to detect ID in any replicate above this line. Note that all panels are in \log_{10} scale.

Figure S6 presents the inbreeding depression (ID) strength estimates (b) for the different inbreeding coefficients (F), with two models in the PEDIGREE populations and with the ADD & DOM & DEMA scenario. The first and second columns depict b estimated with LMM using the weighed (LMM $_{GCTA}^w$) and unweighted (LMM $_{GCTA}^u$) $GCTA$ matrices as random factors, respectively. The first row shows results for the complete PEDIGREE population ($n = 11,924$). The second row shows results for a reduced sample size of the PEDIGREE population ($n = 2,500$, meant to match the size of the 1,000 Genomes Project WORLD population) where subsampled individuals were chosen completely randomly. The third row also shows results for a reduced sample size of the PEDIGREE population ($n = 2,500$) but these individuals were selected to represent the entire spectrum of inbreeding statuses. The violin plots show b estimates distributions among the simulation replicates (100 replicates for the complete population, 10,000 replicates for both subsampled populations). The solid dark grey line is the true strength of ID ($b = -3$). The dashed red line represents the absence of ID ($b = 0$), indicating that we failed to detect ID in any replicate above this line. Root mean square error (RMSE) values associated with both regression models and populations are shown in main table 1. In the complete PEDIGREE population, we see little difference between the three GRMs we tested (figure 1, panel B VS figure S6, panels A and B; table 1): all F yielded accurate estimates of b when used inside a LMM, except for F_{UNI}^u that slightly overestimates the strength of ID while F_{PED} slightly underestimates it. However, when the sample size is reduced to 2,500 individuals, the strength of ID cannot be correctly estimated with F_{PED} , F_{ROH} and F_{HBD} and the LMM $_{GCTA}^u$ model. In addition, only the ranged subsampling allowed correct estimation of b with both $F_{ROH.1MB}$ and $F_{HBD.1MB}$ and the LMM $_{GCTA}^u$ model. This confirms, first than LMM $_{GCTA}^u$ is the least robust among the regression models we used,

and second that F_{ROH} and F_{HBD} based on larger segments are better when studying inbreeding depression.

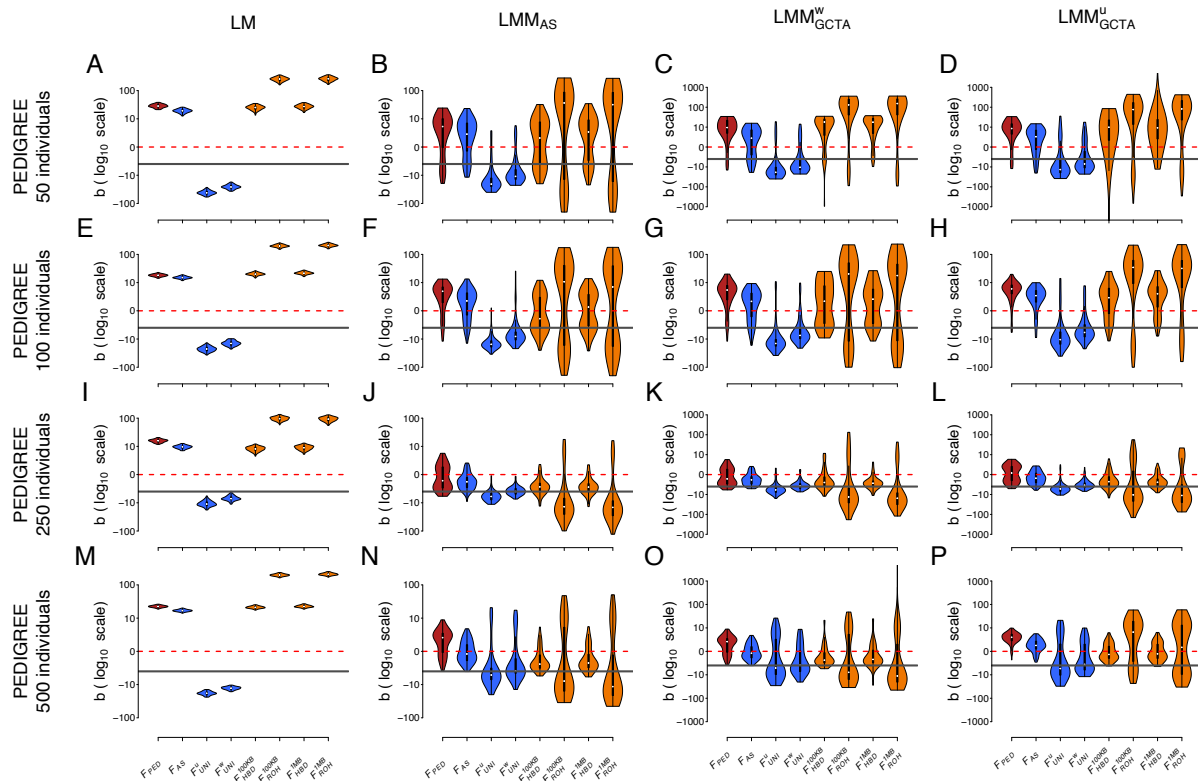


Figure S7: Comparison of the estimation of inbreeding depression strength (b) among different F estimates and models in four simulated PEDIGREE populations of different sizes and with the ADD & DOM & DEMA scenario. Each column represents a regression model. The first column depicts the simple linear regression (panel A, E, I and M), the second column the linear mixed model with allele-sharing GRM matrix as random factor (panel B, F, J and N), the third column the linear mixed model with the $GCTA^w$ relatedness matrix as random factor (panel C, G, K and O) and finally the fourth column represents the linear mixed model with $GCTA^u$ relatedness matrix as random factor (panel D, H, L and P). Inbreeding estimates compared in this figure are F_{PED} , F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$ and $F_{ROH.1MB}$. Each row depicts a simulated PEDIGREE population of different size: 50 individuals in the first row (panels A - D), 100 individuals in the second row (panels E - H), 250 individuals in the third row (panels I - L) and 500 individuals in the fourth row (panels M - P). Violin plots represent the distribution of the inbreeding depression strength estimates (b) among the 100 replicates. The solid dark grey line is the true strength of ID ($b = -3$). The dashed red line represents the absence of ID ($b = 0$), meaning that we failed to detect ID in any replicate above this line. Note that all panels (A - P) are in \log_{10} scale.

Figure S7 presents the inbreeding depression (ID) strength estimates (b) for the different inbreeding coefficients (F), with the four models in smaller PEDIGREE populations and with the ADD & DOM & DEMA scenario. The first column depicts b estimated with the simple LM. The second column depicts LMM_{AS} (using the allele-sharing-based GRM). The third column shows the model using the weighed (LMM_{GCTA}^w) GCTA matrix. The fourth column shows the model using the unweighted (LMM_{GCTA}^u) GCTA matrix. Each

row shows a different population size (50, 100, 250 and 500 individuals). The violin plots show b estimates distributions among the 100 simulation replicates. The solid dark grey line is the true strength of ID ($b = -3$). The dashed red line represents the absence of ID ($b = 0$), indicating that we failed to detect ID in any replicate above this line. For all sample sizes, we see little difference between the three GRMs we tested and F_{UNI}^W always gives the best results. Inbreeding depression is easier to quantify in these small PEDIGREE population because the structure is weaker than the WORLD dataset and the individuals are very inbred.

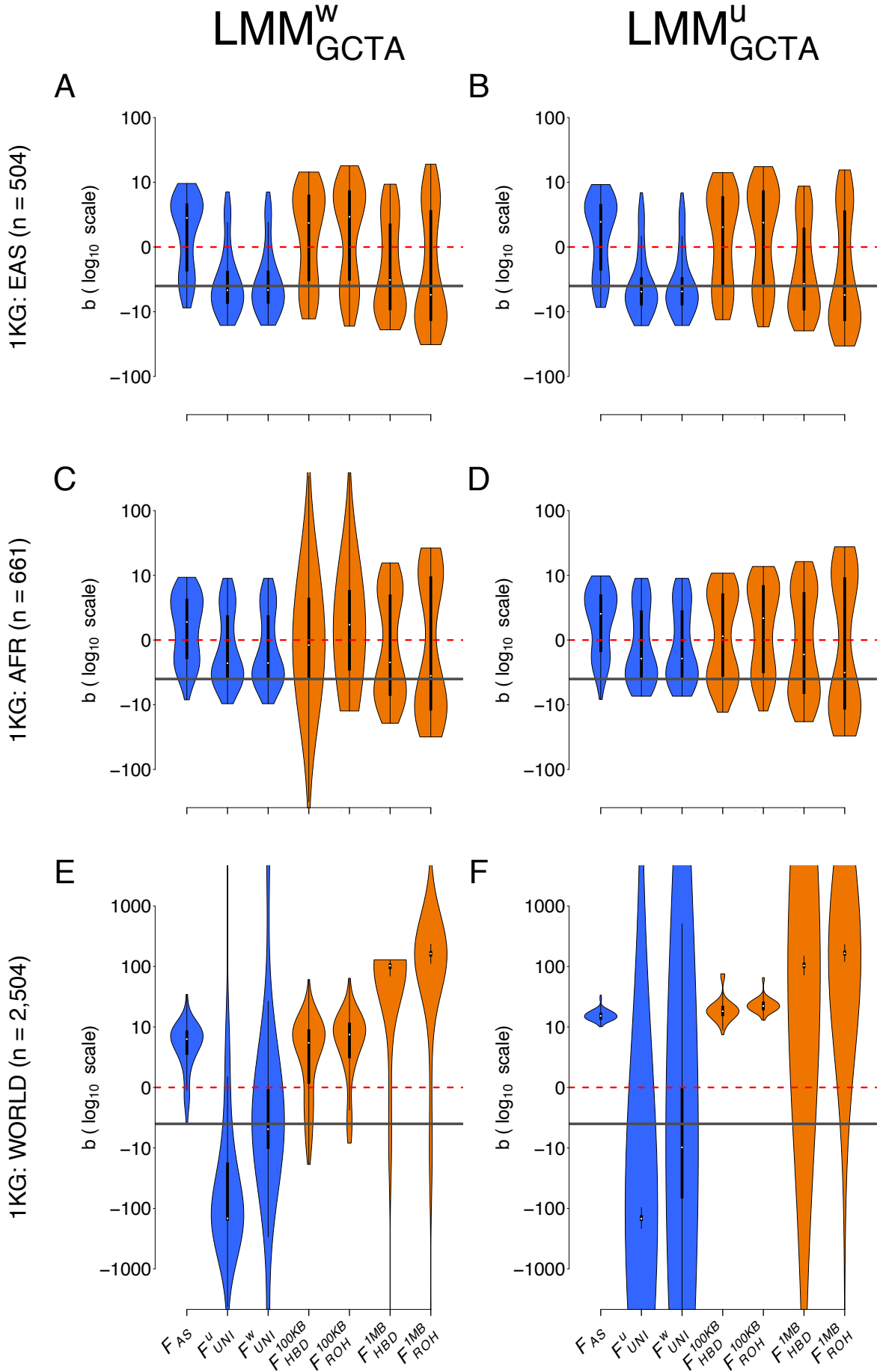


Figure S8: Comparison of the estimation of inbreeding depression strength (b) among different F estimates and the LMM including two different GRMs in the three populations from the 1,000 Genomes Project dataset and with the ADD & DOM & DEMA scenario. The first column depicts the LMM including the $GCTA^w$ matrix (panels A, C and E) and the second column the linear mixed model with the $GCTA^u$ matrix (panels B, D and F). The three rows show the three populations from the 1,000 Genomes project: EAS on panels A and B, AFR on panels C and D and WORLD on panels E and F. Inbreeding estimates presented in this figure are: F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$ and $F_{ROH.1MB}$. Violin plots represent the distribution of the inbreeding depression strength estimates (b) among the 100 simulations replicates. The solid dark grey line is the true strength of ID ($b = -3$). The dashed red line represents the absence of ID ($b = 0$), meaning that we failed to detect ID in any replicate above this line. Note that all panels are in \log_{10} scale.

Figure S8 presents the inbreeding depression (ID) strength estimates (b) for the different inbreeding coefficients (F), with two models in the three populations from the 1,000 Genomes Project: EAS, AFR and WORLD and with the ADD & DOM & DEMA scenario. The first and second columns show b estimated with LMM respectively including the unweighted (LMM_{GCTA^u}) and weighed (LMM_{GCTA^w}) $GCTA$ matrices as random factors. The first row shows results for the EAS population ($n = 504$), the second row shows results for the AFR population ($n = 661$) and the third row shows results for the complete WORLD population ($n = 2,504$). The violin plots show b estimates distributions among the simulation replicates (100 replicates). The solid dark grey line is the true strength of ID ($b = -3$). The dashed red line represents the absence of ID ($b = 0$), indicating that we failed to detect ID in any replicate above this line. RMSE values associated with both regression models and the three populations are shown in main table 2. In the EAS homogeneous population, we see little differences among the three mixed models (figure 2, panel B VS figure S7, panels A and B; table 2). There is also little difference between the three LMM for the three SNPs-based F (i.e. F_{AS} and both F_{UNI}) in the AFR population (panels C and D). However, F_{ROH} and F_{HBD} including smaller segments resulted in larger variance among b estimates with the LMM_{GCTA^w} model (panel C). This was not the case for both LMM_{AS} and LMM_{GCTA^u} models. The variance among b estimates was larger for the AFR population compared to the EAS population. This might be explained by the fact that almost all individuals in the AFR population have a F close to 0. In the EAS population however, the variance in F is larger. The larger RMSE value may also be due to admixture in some AFR individuals for whom F estimation is more complex.

Finally, none of the $GCTA$ -based GRM yielded accurate estimation of b in the highly structured WORLD population. For this last reason, we select the allele-sharing GRM as the best GRM for estimating inbreeding depression.

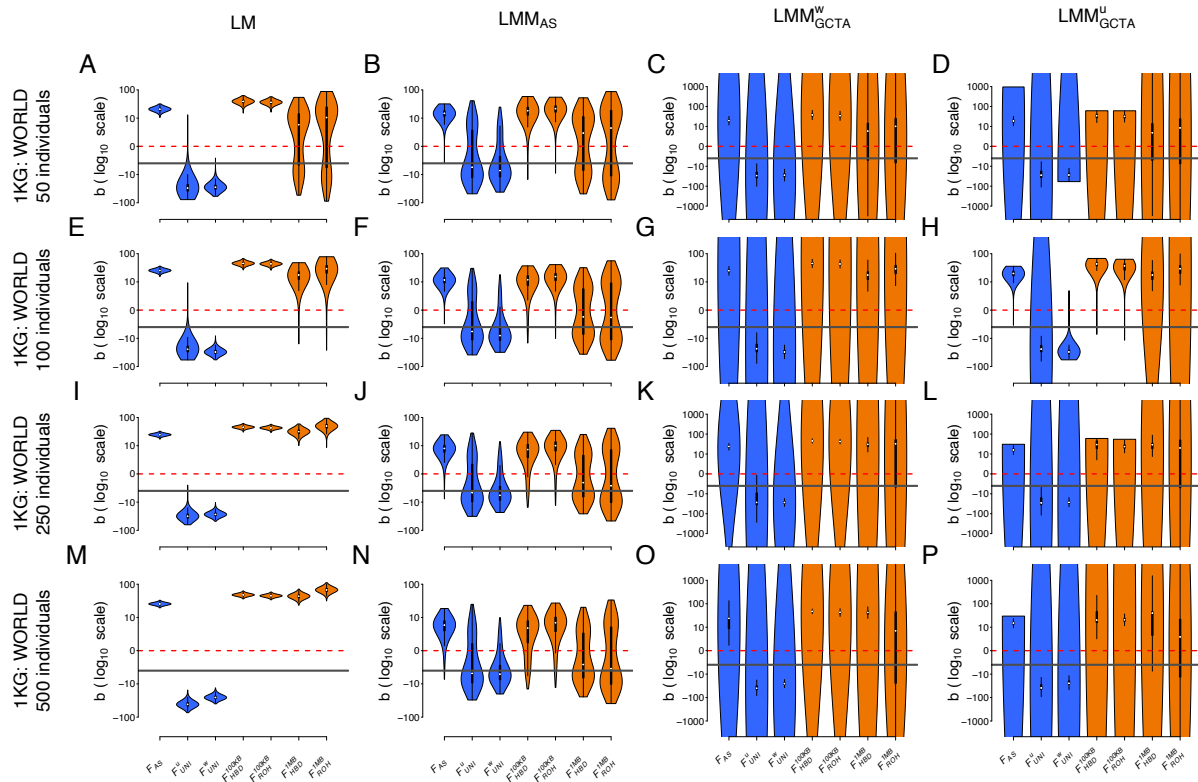


Figure S9: Comparison of the estimation of inbreeding depression strength (b) among different F estimates and models in four subsampling of the WORLD population and with the ADD & DOM & DEMA scenario. Each column represents a regression model. The first column depicts the simple linear regression (panel A, E, I and M), the second column the linear mixed model with allele-sharing GRM matrix as random factor (panel B, F, J and N), the third column the linear mixed model with the $GCTA^w$ relatedness matrix as random factor (panel C, G, K and O) and finally the fourth column represents the linear mixed model with $GCTA^u$ relatedness matrix as random factor (panel D, H, L and P). Inbreeding estimates compared in this figure are F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$ and $F_{ROH.1MB}$. Each row depicts a subsampling of different size: 50 individuals in the first row (panels A - D), 100 individuals in the second row (panels E - H), 250 individuals in the third row (panels I - L) and 500 individuals in the fourth row (panels M - P). Violin plots represent the distribution of the inbreeding depression strength estimates (b) among the 10,000 replicates (100 subsampling replicates and 100 simulation replicates for each subsampling). The solid dark grey line is the true strength of ID ($b = -3$). The dashed red line represents the absence of ID ($b = 0$), meaning that we failed to detect ID in any replicate above this line. Note that all panels (A - P) are in \log_{10} scale.

Figure S9 presents the inbreeding depression (ID) strength estimates (b) for the different inbreeding coefficients (F), with the four models in the smaller subsampled WORLD populations and with the ADD & DOM & DEMA scenario. The first column depicts b estimated with the simple LM. The second column depicts LMM_{AS} (using the allele-sharing-based GRM). The third column shows the model using the weighed (LMM_{GCTA}^w) $GCTA$ matrix. The fourth column shows the model using the unweighted (LMM_{GCTA}^u) $GCTA$ matrix. Each row shows a different population size (50, 100, 250 and 500). The violin plots show b estimates distributions among the 10,000 simulation replicates (100 subsampling replicates and 100 simulations replicates for each subsampling). The solid

dark grey line is the true strength of ID ($b = -3$). The dashed red line represents the absence of ID ($b = 0$), indicating that we failed to detect ID in any replicate above this line. For all sample sizes, we see that the most efficient estimations of b are obtained with the LMM_{AS} model. None of the GCTA-based GRM result in accurate estimation of b . We believe that this difference appears when the structure is strong and when most individuals have low inbreeding coefficients.

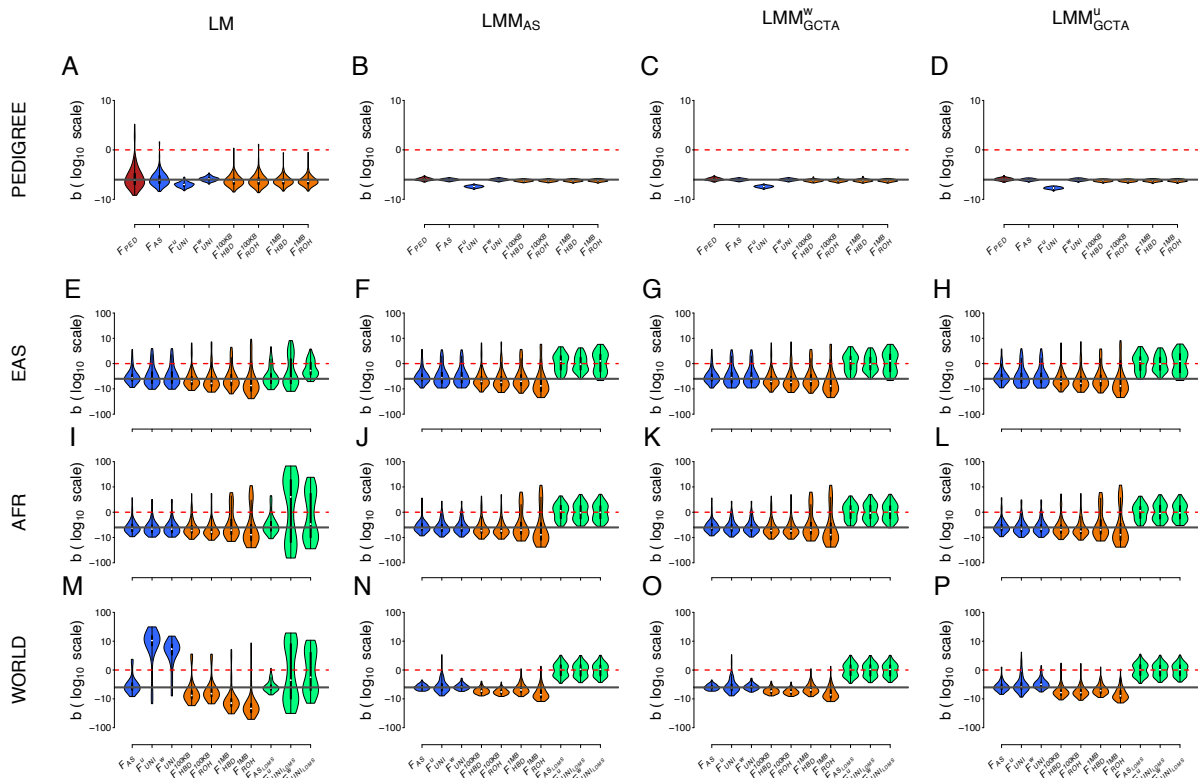


Figure S10: Comparison of the estimation of inbreeding depression strength (b) among different F estimates and models in four different populations with the standard scenario: effect sizes and dominance coefficients randomly assigned to each causal marker and no DEMA. Each column represents a regression model. The first column depicts the simple linear regression (panel A, E, I and M), the second column the linear mixed model with allele sharing GRM matrix as random factor (panel B, F, J and N), the third column the linear mixed model with the $GCTA^w$ relatedness matrix as random factor (panel C, G, K and O) and finally the fourth column represents the linear mixed model with $GCTA^u$ relatedness matrix as random factor (panel D, H, L and P). The first row depicts the complete simulated population (11,924 individuals): PEDIGREE in panels A, B, C and D. Inbreeding estimates compared in these panels (A – D) are F_{PED} , F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$ and $F_{ROH.1MB}$. The last three rows are the populations from the 1,000 Genomes Project: EAS in panels E, F, G and H, AFR in panels I, J, K and L and WORLD in panels M, N, O and P. Inbreeding estimates compared in these panels (E – P) are F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$, $F_{ROH.1MB}$, $F_{AS.LDMS}$, $F_{UNI.LDMS}^u$ and finally $F_{UNI.LDMS}^w$. Violin plots represent the distribution of the inbreeding depression strength estimates (b) among the 100 replicates. The solid dark grey line is the true strength of ID ($b =$

-3). The dashed red line represents the absence of ID ($b = 0$), meaning that we failed to detect ID in any replicate above this line. Note that all panels (A - P) are in \log_{10} scale.

Figure S10 presents the results of ID strength estimation for the standard scenario (additive effect sizes and dominance coefficients are randomly drawn and there is no DEMA). Corresponding RMSE values can be found in tables S3 – S6. In the complete pedigree population, we can see that all regression models (but especially the mixed models) and all inbreeding coefficients result in efficient estimates (we use efficient to describe an estimate with low RMSE, thus which is unbiased and has low variance) (panels A – D, tables S4 – S6). The most efficient estimate was F_{UNI}^W . The variance among b estimates was larger for the three populations of the 1,000 Genomes Project and some replicates resulted in estimated b above 0 (panels E – P). This is due to the smaller sample sizes. In both the EAS and AFR populations, F_{AS} , both F_{UNI} and F_{ROH} gave unbiased estimates for all models (panels E – L). However, LDMS-based F were always biased, especially for all the mixed models where they were centered around 0 (panels E – L, tables S4 – S6). We believe it is because the sample sizes we used were too small to correctly estimate allelic frequencies per MAF and LD bins. With this scenario, F_{AS} (and F_{ASLDMS} in the simple LM model) gives the most efficient estimate of ID in the WORLD population (panels E – P, tables S4 – S6). In addition, the three LMM perform similarly.

To conclude, when additive and dominance effects are uniform, and there is no DEMA, F_{AS} results in efficient estimates of ID (even with a simple linear model and strong population structure: panel M). With a mixed model, it is possible to estimate ID correctly for all inbreeding coefficients, except for LDMS-based F .

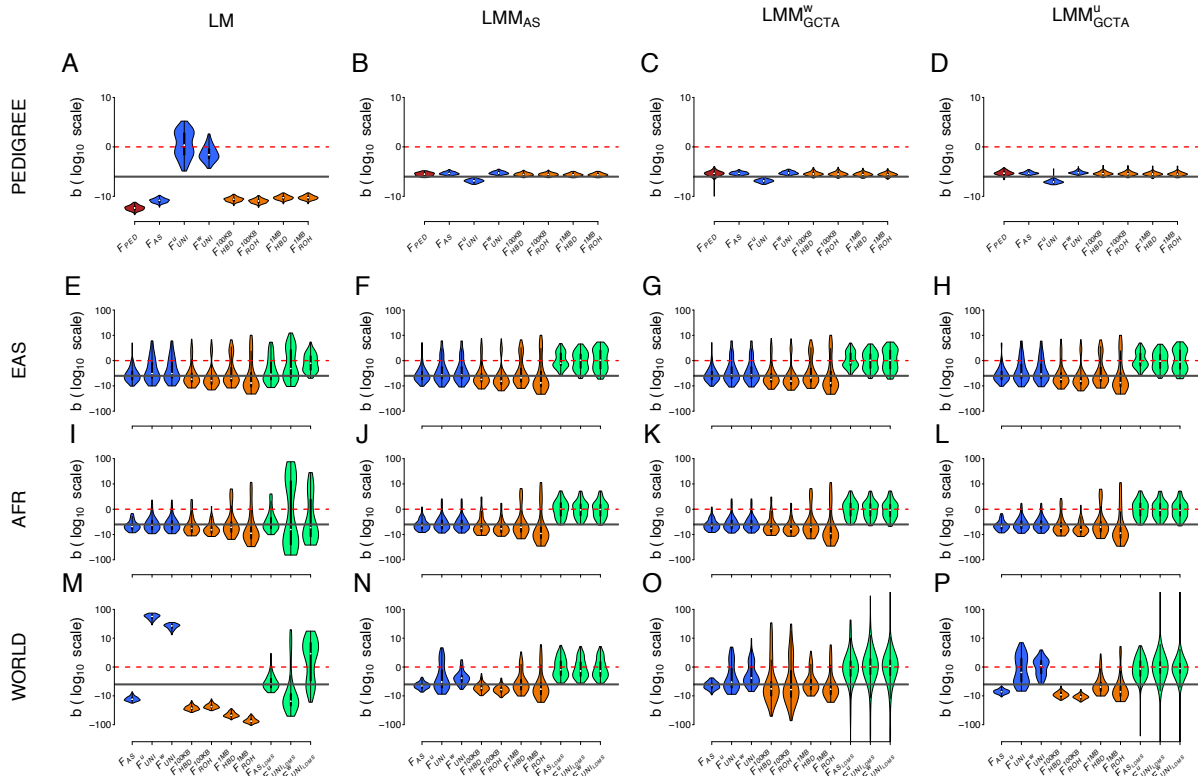


Figure S11: Comparison of the estimation of inbreeding depression strength (b) among different F estimates and models in four different populations with the ADD scenario: effect sizes assigned proportional to the MAF of causal markers, dominance coefficients randomly assigned to causal markers and no DEMA. Each column represents a regression model. The first column depicts the simple linear regression (panel A, E, I and M), the second column the linear mixed model with allele sharing GRM matrix as random factor (panel B, F, J and N), the third column the linear mixed model with the $GCTA^w$ relatedness matrix as random factor (panel C, G, K and O) and finally the fourth column represents the linear mixed model with $GCTA^u$ relatedness matrix as random factor (panel D, H, L and P). The first row depicts the complete simulated population (11,924 individuals): PEDIGREE in panels A, B, C and D. Inbreeding estimates compared in these panels (A – D) are F_{PED} , F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$ and $F_{ROH.1MB}$. The last three rows are the populations from the 1,000 Genomes Project: EAS in panels E, F, G and H, AFR in panels I, J, K and L and WORLD in panels M, N, O and P. Inbreeding estimates compared in these panels (E – P) are F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$, $F_{ROH.1MB}$, $F_{AS.LDMS}$, $F_{UNI.LDMS}^u$ and finally $F_{UNI.LDMS}^w$. Violin plots represent the distribution of the inbreeding depression strength estimates (b) among the 100 replicates. The solid dark grey line is the true strength of ID ($b = -3$). The dashed red line represents the absence of ID ($b = 0$), meaning that we failed to detect ID in any replicate above this line. Note that all panels (A – P) are in \log_{10} scale. Also note that linear mixed models did not converge for some replicates (yielding estimated b values above 1,000 or below -1,000, not shown if outside the graph limits). Percentages of replicates which did not converge: panel O (WORLD, $GCTA^w$): 1%, for $F_{AS.LDMS}$, 3% for $F_{UNI.LDMS}^u$ and 1% for $F_{UNI.LDMS}^w$; panel P (WORLD, $GCTA^u$): 3% for $F_{UNI.LDMS}^u$ and 2% for $F_{UNI.LDMS}^w$.

Figure S11 presents the results of ID strength estimation for the ADD scenario (when the additive effects are inversely proportional to MAF, the dominance effects are independent of MAF and there is no DEMA). Corresponding RMSE values can be found in tables S3 – S6. The simple LM results in biased b estimates with an overestimation of ID for all

inbreeding coefficients except both F_{UNI} , which underestimate it (panel A, tables S4 – S6). The three mixed models, on the other hand, provide efficient estimates of inbreeding depression in the PEDIGREE population (panels B – D, tables S4 – S6). In the EAS population, all but the LDMS-based F are unbiased, but with a large variance and we see no improvement of using a LMM rather than a simple LM (panels E – H, tables S4 – S6). Interestingly, we see similar results for the AFR population but with lower variance (panels I – L, tables S4 – S6). All estimates in the WORLD population, however, are strongly biased with the simple LM: ID is underestimated with F_{AS} and all F_{ROH} and overestimated with both F_{UNI} (panel M, tables S4 – S6). For the WORLD population and the three LMM, F_{AS} yields the most efficient estimation of b , especially in combination with the LMM_{AS} model (for which the variance around b estimates is the smallest) (panels N – P, tables S4 – S6).

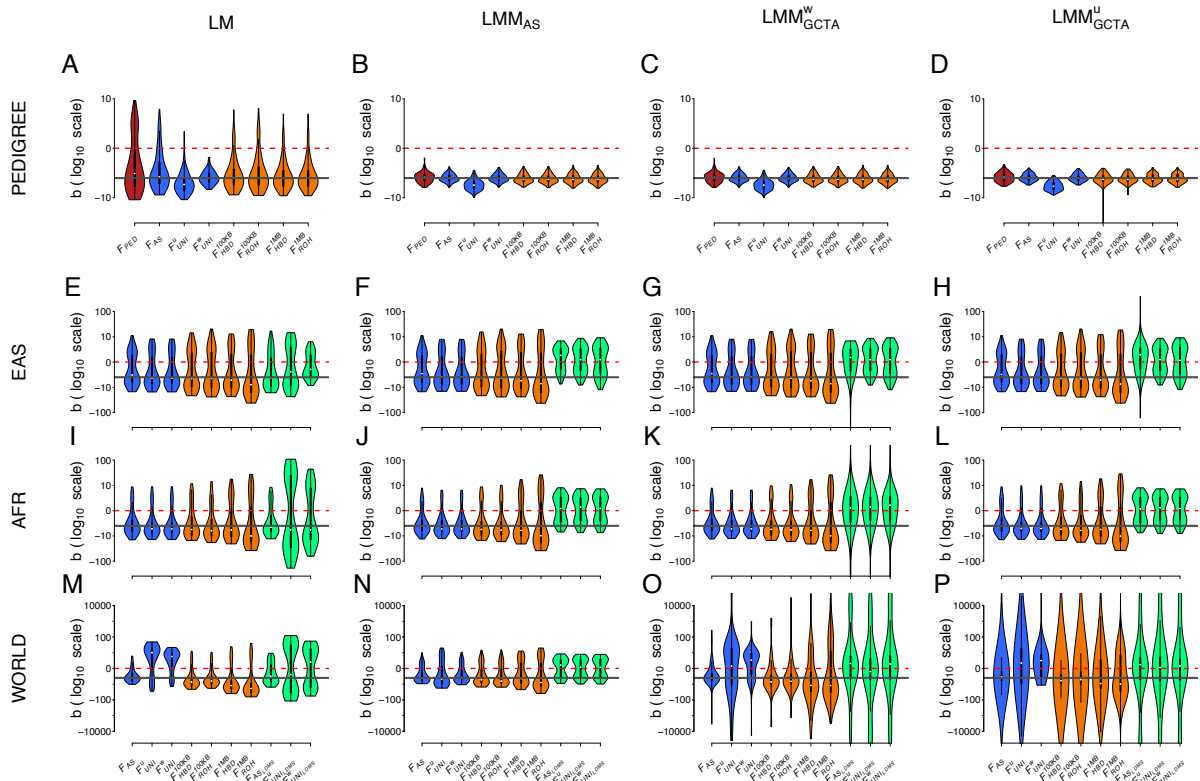


Figure S12: Comparison of the estimation of inbreeding depression strength (b) among different F estimates and models in four different populations with the DOM scenario: effect sizes randomly assigned to causal markers, dominance coefficients assigned proportional to the MAF of causal markers and no DEMA. Each column represents a regression model. The first column depicts the simple linear regression (panel A, E, I and M), the second column the linear mixed model with allele sharing GRM matrix as random factor (panel B, F, J and N), the third column the linear mixed model with the GCTA^w relatedness matrix as random factor (panel C, G, K and O) and finally the fourth column represents the linear mixed model with GCTA^u relatedness matrix as random factor (panel D, H, L and P). The first row depicts the complete simulated population (11,924 individuals): PEDIGREE in panels A, B, C and D. Inbreeding estimates compared in

these panels (A – D) are F_{PED} , F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$ and $F_{ROH.1MB}$. The last three rows are the populations from the 1,000 Genomes Project: EAS in panels E, F, G and H, AFR in panels I, J, K and L and WORLD in panels M, N, O and P. Inbreeding estimates compared in these panels (E – P) are F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$, $F_{ROH.1MB}$, $F_{AS.LDMS}$, $F_{UNI.LDMS}^u$ and finally $F_{UNI.LDMS}^w$. Violin plots represent the distribution of the inbreeding depression strength estimates (b) among the 100 replicates. The solid dark grey line is the true strength of ID ($b = -3$). The dashed red line represents the absence of ID ($b = 0$), meaning that we failed to detect ID in any replicate above this line. Note that all panels (A – P) are in \log_{10} scale. Also note that linear mixed models did not converge for some replicates (yielding estimated b values above 1,000 or below -1,000, not shown if outside the graph limits). Percentages of replicates which did not converge: panel K (AFR, $GCTA^w$): 8%, for $F_{AS.LDMS}$, 7% for $F_{UNI.LDMS}^u$ and 4% for $F_{UNI.LDMS}^w$; panel O (WORLD, $GCTA^w$): 1%, for F_{AS} , 10% for F_{UNI}^u , 1% for F_{UNI}^w , for, 6% for $F_{HBD.100KB}$, 7% for $F_{ROH.100KB}$, 8% for $F_{HBD.1MB}$, 4% $F_{ROH.1MB}$, 22% for $F_{AS.LDMS}$, 31% for $F_{UNI.LDMS}^u$ and 31% for $F_{UNI.LDMS}^w$; panel P (WORLD, $GCTA^u$) 10%, for F_{AS} , 9% for F_{UNI}^u , 2% for F_{UNI}^w , 7% for $F_{HBD.100KB}$, 7% for $F_{ROH.100KB}$, 5% for $F_{HBD.1MB}$, 2% $F_{ROH.1MB}$, 22% for $F_{AS.LDMS}$, 31% for $F_{UNI.LDMS}^u$ and 28% for $F_{UNI.LDMS}^w$.

Figure S12 shows the DOM scenario (where dominance effects are inversely proportional to MAF while additive effects are independent and there is no DEMA). Corresponding RMSE values can be found in tables S3 – S6. In the PEDIGREE population and with the simple LM, only F_{UNI}^w gives efficient estimation of b ; the other inbreeding coefficients are unbiased, but with large variance (panel A, table S3). For the same PEDIGREE population, the three mixed models reduce the variance around b estimates (panels B – D, tables S4 – S6). In the EAS and AFR populations, the results are similar to what we observed in figure S11: all the inbreeding coefficients are mostly unbiased but exhibit large variance (especially in the EAS population, panels E – L, tables S3 – S6). Concerning the WORLD population, the most efficient estimates of ID are obtained with F_{AS} (except for the LMM_{GCTA}^u model, panels M – P, tables S3 – S6). With both $GCTA$ GRM matrices, all F results in large variance around b estimates (panels O and P, tables S5 and S6).

To conclude, LMM_{AS} correctly estimates ID with the DOM scenario, in particular with F_{AS} . However, we observe very large variance with smaller sample sizes: with some replicates overlapping with 0 in the WORLD population and many replicates in both the EAS and AFR populations. This is probably because the sample sizes are too small to estimate ID in these two populations.

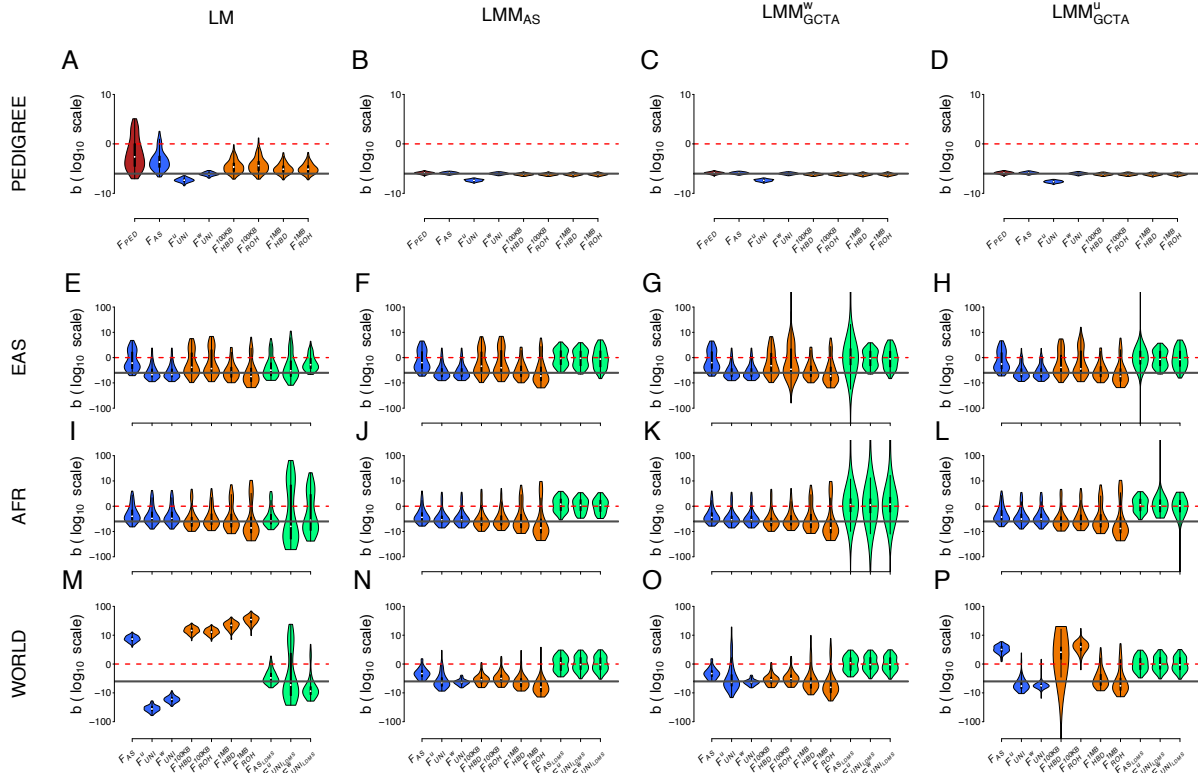


Figure S13: Comparison of the estimation of inbreeding depression strength (b) among different F estimates and models in four different populations with the DEMA scenario: effect sizes and dominance coefficients randomly assigned to causal markers and presence of DEMA. Each column represents a regression model. The first column depicts the simple linear regression (panel A, E, I and M), the second column the linear mixed model with allele sharing GRM matrix as random factor (panel B, F, J and N), the third column the linear mixed model with the $GCTA^w$ relatedness matrix as random factor (panel C, G, K and O) and finally the fourth column represents the linear mixed model with $GCTA^u$ relatedness matrix as random factor (panel D, H, L and P). The first row depicts the complete simulated population (11,924 individuals): PEDIGREE in panels A, B, C and D. Inbreeding estimates compared in these panels (A – D) are F_{PED} , F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$ and $F_{ROH.1MB}$. The last three rows are the populations from the 1,000 Genomes Project: EAS in panels E, F, G and H, AFR in panels I, J, K and L and WORLD in panels M, N, O and P. Inbreeding estimates compared in these panels (E – P) are F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$, $F_{ROH.1MB}$, $F_{AS.LDMS}$, $F_{UNI.LDMS}^u$ and finally $F_{UNI.LDMS}^w$. Violin plots represent the distribution of the inbreeding depression strength estimates (b) among the 100 replicates. The solid dark grey line is the true strength of ID ($b = -3$). The dashed red line represents the absence of ID ($b = 0$), meaning that we failed to detect ID in any replicate above this line. Note that all panels (A – P) are in \log_{10} scale. Also note that linear mixed models did not converge for some replicates (yielding estimated b values above 1,000 or below -1,000, not shown if outside the graph limits). Percentages of replicates which did not converge: panel G (EAS, $GCTA^w$): 1% for $F_{ROH.100KB}$, 12% for $F_{AS.LDMS}$; panel H (EAS, $GCTA^u$): 2% for $F_{AS.LDMS}$; panel K (AFR, $GCTA^w$): 11%, for $F_{AS.LDMS}$, 17% for $F_{UNI.LDMS}^u$ and 14% for $F_{UNI.LDMS}^w$; panel L (AFR, $GCTA^u$): 1% for $F_{UNI.LDMS}^u$ and 1% for $F_{UNI.LDMS}^w$; panel P (WORLD, $GCTA^u$) 1% for $F_{HBD.100KB}$.

Figure S13 shows the DEMA scenario (where both additive effects and dominance coefficients are independent of MAF but there is DEMA). Corresponding RMSE values can be found in tables S3 – S6. With the simple LM and in the large PEDIGREE population, only FUNI and to a lesser extent IDB segments-based F estimated with larger

segments ($F_{HBD.1MB}$ and $F_{ROH.1MB}$) yield unbiased estimation of b (panel A, table S2). Meanwhile, in the same large PEDIGREE population, the three mixed models allow efficient estimation of b with all F except F_{UNI}^u , which overestimated the strength of ID (panels B – D, tables S4 – S6). In the three populations from the 1,000 Genomes Project and especially in the EAS and WORLD populations, F_{AS} is constantly overestimating the strength of ID (panels E – P, tables S3 – S6). This is because DEMA is included in the simulations; F_{AS} strongly correlates with the minor allele count (MAC) which makes it sensitive to DEMA [56]. Interestingly, LDMS stratification, allowed to get rid of the bias introduced by DEMA on F_{AS} but it did not result in the most efficient estimation of b (panel M, table S3 – S6). Indeed, the lowest RMSE values are obtained with F_{UNI}^w , and both the LMM_{AS} and LMM_{GCTA}^w models confirming that F_{UNI}^w is robust to DEMA (tables S3 – S6). With both LMM_{AS} and LMM_{GCTA}^w models, IDB segments-based F estimated with larger segments ($F_{HBD.1MB}$ and $F_{ROH.1MB}$) are just behind F_{UNI}^w in term of efficiency (tables S3 – S6).

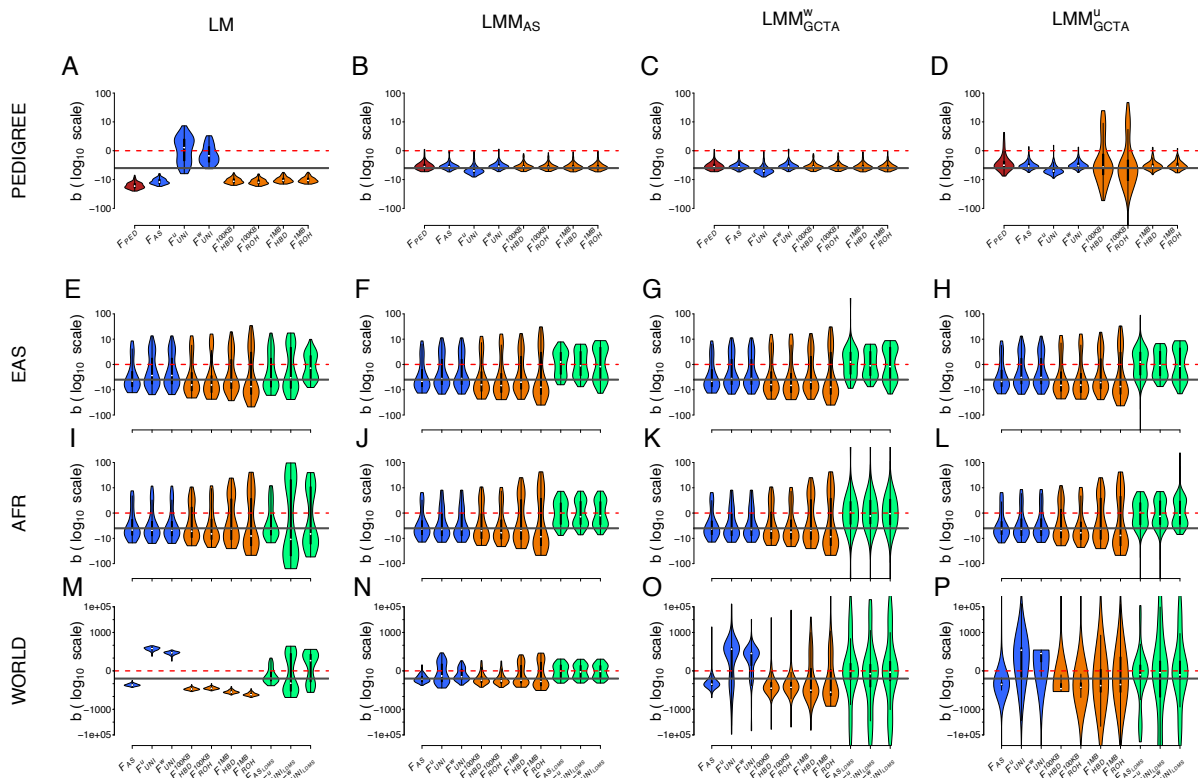


Figure S14: Comparison of the estimation of inbreeding depression strength (b) among different F estimates and models in four different populations with the ADD & DOM scenario: effect sizes and dominance coefficients assigned proportional to the MAF of causal markers and no DEMA. Each column represents a regression model. The first column depicts the simple linear regression (panel A, E, I and M), the second column the linear mixed model with allele sharing GRM matrix as random factor (panel B, F, J and N), the third column the linear mixed model with the $GCTA^w$ relatedness matrix as random factor (panel C, G, K

and O) and finally the fourth column represents the linear mixed model with $GCTA^u$ relatedness matrix as random factor (panel D, H, L and P). The first row depicts the complete simulated population (11,924 individuals): PEDIGREE in panels A, B, C and D. Inbreeding estimates compared in these panels (A – D) are F_{PED} , F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$ and $F_{ROH.1MB}$. The last three rows are the populations from the 1,000 Genomes Project: EAS in panels E, F, G and H, AFR in panels I, J, K and L and WORLD in panels M, N, O and P. Inbreeding estimates compared in these panels (E – P) are F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$, $F_{ROH.1MB}$, $F_{AS.LDMS}$, $F_{UNI.LDMS}^u$ and finally $F_{UNI.LDMS}^w$. Violin plots represent the distribution of the inbreeding depression strength estimates (b) among the 100 replicates. The solid dark grey line is the true strength of ID ($b = -3$). The dashed red line represents the absence of ID ($b = 0$), meaning that we failed to detect ID in any replicate above this line. Note that all panels (A – P) are in \log_{10} scale. Also note that linear mixed models did not converge for some replicates (yielding estimated b values above 1,000 or below -1,000, not shown if outside the graph limits). Percentages of replicates which did not converge: panel D (PEDIGREE, $GCTA^u$): 1% for $F_{ROH.100KB}$; panel G (EAS, $GCTA^w$): 1% for $F_{AS.LDMS}$; panel H (EAS, $GCTA^u$): 1% for $F_{AS.LDMS}$; panel K (AFR, $GCTA^w$): 5%, for $F_{AS.LDMS}$, 5% for $F_{UNI.LDMS}^u$ and 7% for $F_{UNI.LDMS}^w$; panel L (AFR, $GCTA^u$): 1% for $F_{AS.LDMS}$ and 1% for $F_{UNI.LDMS}^u$; panel O (WORLD, $GCTA^w$): 1% for F_{AS} , 5% F_{UNI}^u and 3% F_{UNI}^w , 5% for $F_{HBD.100KB}$, 8% for $F_{ROH.100KB}$, 6% for $F_{HBD.1MB}$, 3% for $F_{ROH.1MB}$, 28% for $F_{AS.LDMS}$, 32% for $F_{UNI.LDMS}^u$ and finally 33% for $F_{UNI.LDMS}^w$; panel P (WORLD, $GCTA^u$) 4% for F_{AS} , 9% F_{UNI}^u , 4% F_{UNI}^w , 3% for $F_{HBD.100KB}$, 6% for $F_{ROH.100KB}$, 11% for $F_{HBD.1MB}$, 11% for $F_{ROH.1MB}$, 22% for $F_{AS.LDMS}$, 28% for $F_{UNI.LDMS}^u$ and finally 32% for $F_{UNI.LDMS}^w$.

Figure S14 shows the ADD & DOM scenario (where both additive effects and dominance coefficients are proportional to MAF and there is no DEMA). Corresponding RMSE values can be found in tables S3 – S6. In the complete PEDIGREE population and with the simple LM, the strength of ID is underestimated by both FUNI and overestimated by all other F (panel A, table S3). Similarly to what was observed with the previous scenarios, both LMM_{AS} and LMM_{GCTA^w} models allow efficient estimation of b with all the F (panels B and C, tables S4 – S5). However, the LMM_{GCTA^u} model results in biased estimation of b with the short IBD segments-based F ($F_{HBD.100KB}$ and $F_{ROH.100KB}$, panel D, table S5). Similarly to what was observed before, we see no difference between the four models (except for the LDMS-based F) in both the EAS and AFR homogeneous populations (panels E – L, tables S3 – S6). In the highly structured WORLD population however, the lowest RMSE values are obtained with the LMM_{AS} model and especially with F_{AS} (closely followed by F_{UNI}^w panel N, tables S3 – S6). This is because DEMA is not included in this model and would strongly bias b estimation with F_{AS} . RMSE values are much larger for all F with both $GCTA$ matrices (panels O and P, tables S3 – S6).

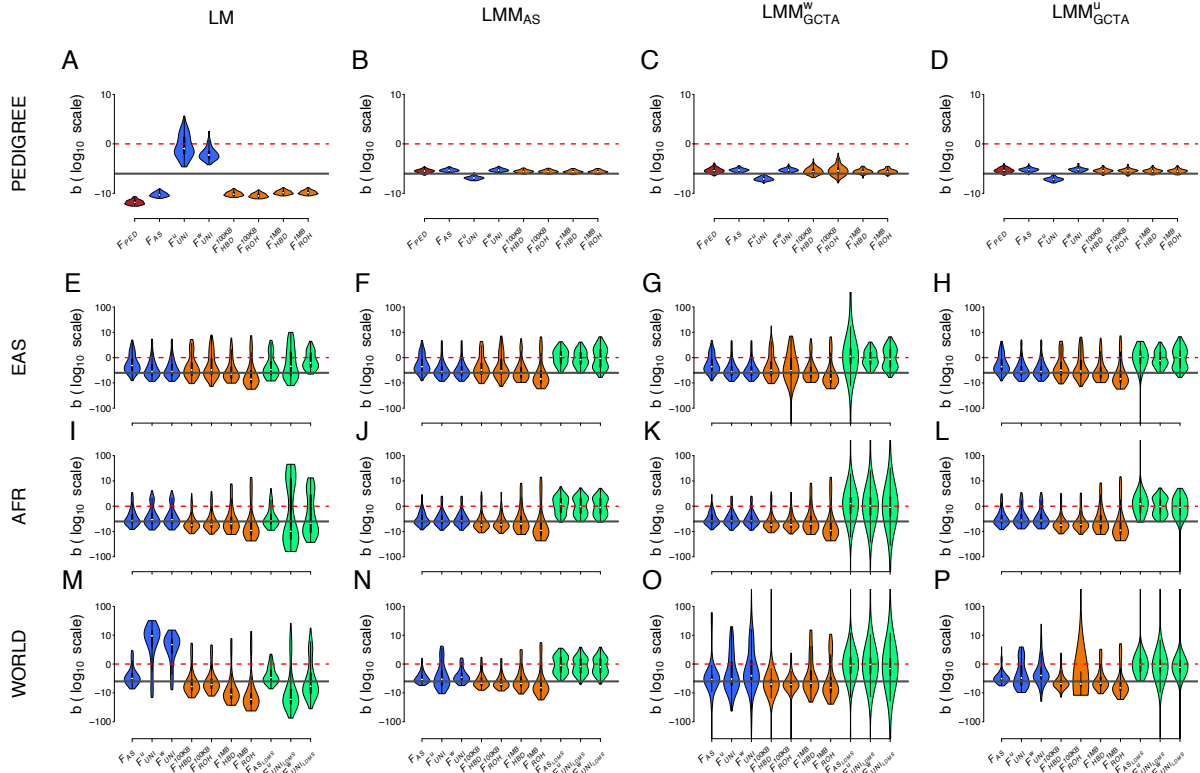


Figure S15: Comparison of the estimation of inbreeding depression strength (b) among different F estimates and models in four different populations with the ADD & DEMA scenario: effect sizes assigned proportional to the MAF of causal markers, dominance coefficients randomly assigned to causal markers and presence of DEMA. Each column represents a regression model. The first column depicts the simple linear regression (panel A, E, I and M), the second column the linear mixed model with allele-sharing GRM matrix as random factor (panel B, F, J and N), the third column the linear mixed model with the $GCTA^w$ relatedness matrix as random factor (panel C, G, K and O) and finally the fourth column represents the linear mixed model with $GCTA^u$ relatedness matrix as random factor (panel D, H, L and P). The first row depicts the complete simulated population (11,924 individuals): PEDIGREE in panels A, B, C and D. Inbreeding estimates compared in these panels (A – D) are F_{PED} , F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$ and $F_{ROH.1MB}$. The last three rows are the populations from the 1,000 Genomes Project: EAS in panels E, F, G and H, AFR in panels I, J, K and L and WORLD in panels M, N, O and P. Inbreeding estimates compared in these panels (E – P) are F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$, $F_{ROH.1MB}$, $F_{AS.LDMS}$, $F_{UNI.LDMS}^u$ and finally $F_{UNI.LDMS}^w$. Violin plots represent the distribution of the inbreeding depression strength estimates (b) among the 100 replicates. The solid dark grey line is the true strength of ID ($b = -3$). The dashed red line represents the absence of ID ($b = 0$), meaning that we failed to detect ID in any replicate above this line. Note that all panels (A – P) are in \log_{10} scale. Also note that linear mixed models did not converge for some replicates (yielding estimated b values above 1,000 or below -1,000. not shown if outside the graph limits). Percentages of replicates which did not converge: panel G (EAS, $GCTA^w$): 1% for $F_{ROH.100KB}$, and 16% for $F_{AS.LDMS}$; panel H (EAS, $GCTA^u$): 2% for $F_{AS.LDMS}$; panel K (AFR, $GCTA^w$): 13%, for $F_{AS.LDMS}$, 13% for $F_{UNI.LDMS}^u$ and 15% for $F_{UNI.LDMS}^w$; panel L (AFR, $GCTA^u$): 2% for $F_{AS.LDMS}$ and 1% for $F_{UNI.LDMS}^u$; panel O (WORLD, $GCTA^w$): 1% for F_{AS} , 2% F_{UNI}^w , 3% for $F_{HBD.100KB}$, 1% for $F_{ROH.100KB}$, 7% for $F_{AS.LDMS}$, 14% for $F_{UNI.LDMS}^u$ and finally 19% for $F_{UNI.LDMS}^w$; panel P (WORLD, $GCTA^u$) 1% for $F_{ROH.100KB}$, 2% for $F_{AS.LDMS}$, 4% for $F_{UNI.LDMS}^u$ and finally 2% for $F_{UNI.LDMS}^w$.

Figure S15 shows the results of ID strength estimation for the ADD & DEMA scenario (when the additive effects are inversely proportional to MAF, the dominance effects are

independent of MAF and there is presence of DEMA). Corresponding RMSE values can be found in tables S3 – S6. Similarly to what was observed in the previous figures with the simple LM, the strength of ID is underestimated by both F_{UNI} and overestimated by all other F in the complete PEDIGREE population (panel A, table S3). The LMM_{AS} and LMM_{GCTA^U} models allow for an efficient estimation of b with all F except F_{UNI}^u which results in a lightly overestimated estimation of b (panels B and D, tables S3 – S6). As for the LMM_{GCTA^W} model, F_{UNI}^u overestimates the strength of ID and the variance among b estimates was larger with both short IBD segments-based F ($F_{HBD.100KB}$ and $F_{ROH.100KB}$, panel C, table S6). In both the homogeneous EAS and AFR populations there are not much differences among the models (panels E – L, tables S3 – S6). However, in the WORLD population, the smallest RMSE are obtained with the LMM_{AS} model (panel N, tables S3 – S6).

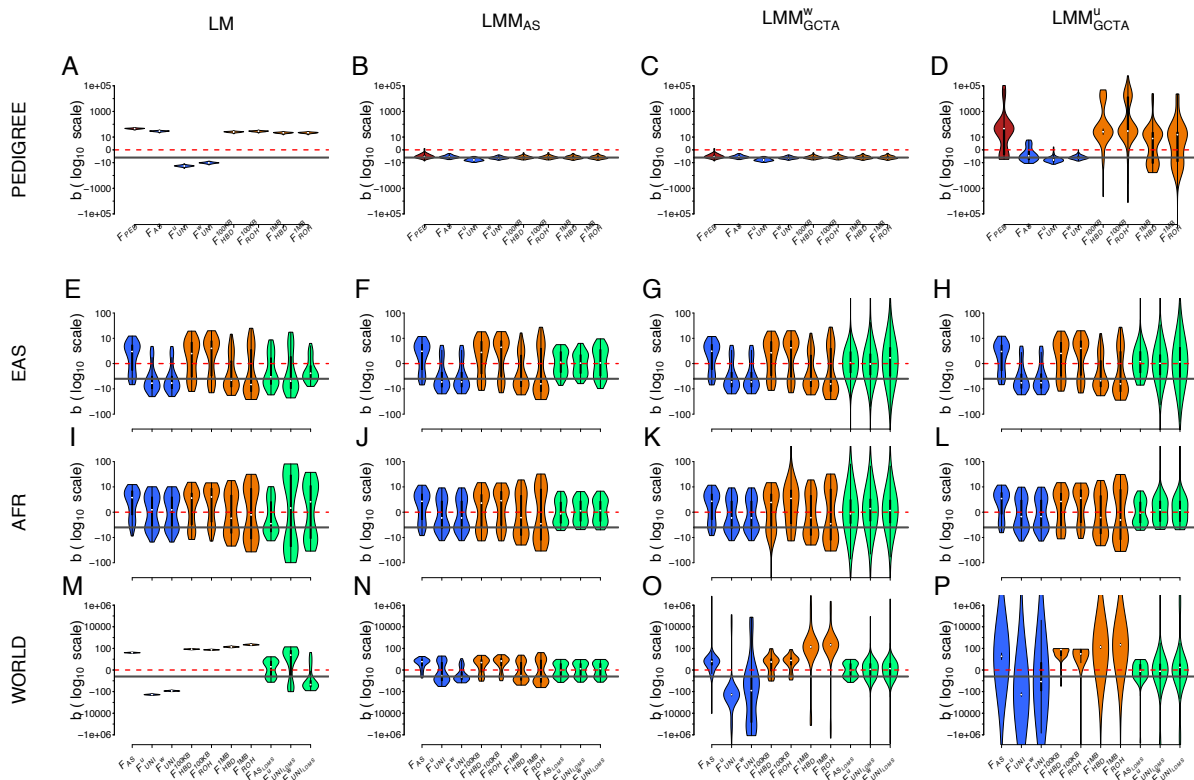


Figure S16: Comparison of the estimation of inbreeding depression strength (b) among different F estimates and models in four different populations with the DOM & DEMA scenario: effect sizes randomly assigned to causal markers, dominance coefficients assigned proportional to the MAF of causal markers and presence of DEMA. Each column represents a regression model. The first column depicts the simple linear regression (panel A, E, I and M), the second column the linear mixed model with allele-sharing GRM matrix as random factor (panel B, F, J and N), the third column the linear mixed model with the $GCTA^W$ relatedness matrix as random factor (panel C, G, K and O) and finally the fourth column represents the linear mixed model with $GCTA^U$ relatedness matrix as random factor (panel D, H, L and P). The first row depicts the complete simulated population (11,924 individuals): PEDIGREE in panels A, B, C and D. Inbreeding estimates

compared in these panels (A – D) are F_{PED} , F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$ and $F_{ROH.1MB}$. The last three rows are the populations from the 1,000 Genomes Project: EAS in panels E, F, G and H, AFR in panels I, J, K and L and WORLD in panels M, N, O and P. Inbreeding estimates compared in these panels (E – P) are F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$, $F_{ROH.1MB}$, $F_{AS.LDMS}$, $F_{UNI.LDMS}^u$ and finally $F_{UNI.LDMS}^w$. Violin plots represent the distribution of the inbreeding depression strength estimates (b) among the 100 replicates. The solid dark grey line is the true strength of ID ($b = -3$). The dashed red line represents the absence of ID ($b = 0$), meaning that we failed to detect ID in any replicate above this line. Note that all panels (A – P) are in \log_{10} scale. Also note that linear mixed models did not converge for some replicates (yielding estimated b values above 1,000 or below -1,000, not shown if outside the graph limits). Percentages of replicates which did not converge: panel D (PEDIGREE, $GCTA^u$): 10% for F_{PED} , 25% for $F_{HBD.100KB}$, 31% for $F_{ROH.100KB}$, 4% for $F_{HBD.1MB}$, 9% for $F_{ROH.1MB}$; panel G (EAS, $GCTA^w$): 3% for $F_{AS.LDMS}$, 8% for $F_{UNI.LDMS}^u$ and 8% for $F_{UNI.LDMS}^w$; panel H (EAS, $GCTA^u$): 3% for $F_{AS.LDMS}$, 10% for $F_{UNI.LDMS}^u$ and 10% for $F_{UNI.LDMS}^w$; panel K (AFR, $GCTA^w$): 1% for $F_{HBD.100KB}$, 2% for $F_{ROH.100KB}$, 15% for $F_{AS.LDMS}$, 9% for $F_{UNI.LDMS}^u$ and 12% for $F_{UNI.LDMS}^w$; panel L (AFR, $GCTA^u$): 1% for $F_{UNI.LDMS}^u$ and 1% for $F_{UNI.LDMS}^w$; panel O (WORLD, $GCTA^w$): 4% for F_{AS} , 15% F_{UNI}^u , 45% F_{UNI}^w , 16% for $F_{HBD.1MB}$, 12% for $F_{ROH.1MB}$, 3% for $F_{UNI.LDMS}^u$ and 3% for $F_{UNI.LDMS}^w$; panel P (WORLD, $GCTA^u$) 15% for F_{AS} , 28% F_{UNI}^u , 26% F_{UNI}^w , 1% for $F_{ROH.100KB}$, 14% for $F_{ROH.1MB}$, 10% for $F_{HBD.1MB}$, 2% for $F_{AS.LDMS}$, 13% for $F_{UNI.LDMS}^u$ and finally 16% for $F_{UNI.LDMS}^w$.

Figure S16 shows the results of ID strength estimation for the DOM & DEMA scenario (when the additive effects are randomly assigned, the dominance effects are inversely proportional to MAF and there is presence of DEMA). Corresponding RMSE values can be found in tables S3 – S6. Interestingly, we find the opposite of previous findings with LM in the entire PEDIGREE population: both F_{UNI} overestimate the strength of ID whereas all the other F underestimate it (panel A). Both LMM_{AS} and LMM_{GCTA}^w models allow for efficient estimation of b with all the F (panels B and C, tables S4 – S5). With the LMM_{GCTA}^u however, the estimation of b for F_{PED} and all IBD segments-based F was really poor. There is no difference between the four regression models for both homogeneous populations (panels E – L, tables S3 – S6). Concerning the WORLD population, the smallest RMSE values were only obtained with the LMM_{AS} model, both $GCTA$ matrices resulted in very large variances among b estimates (panels M – P, tables S3 – S6).

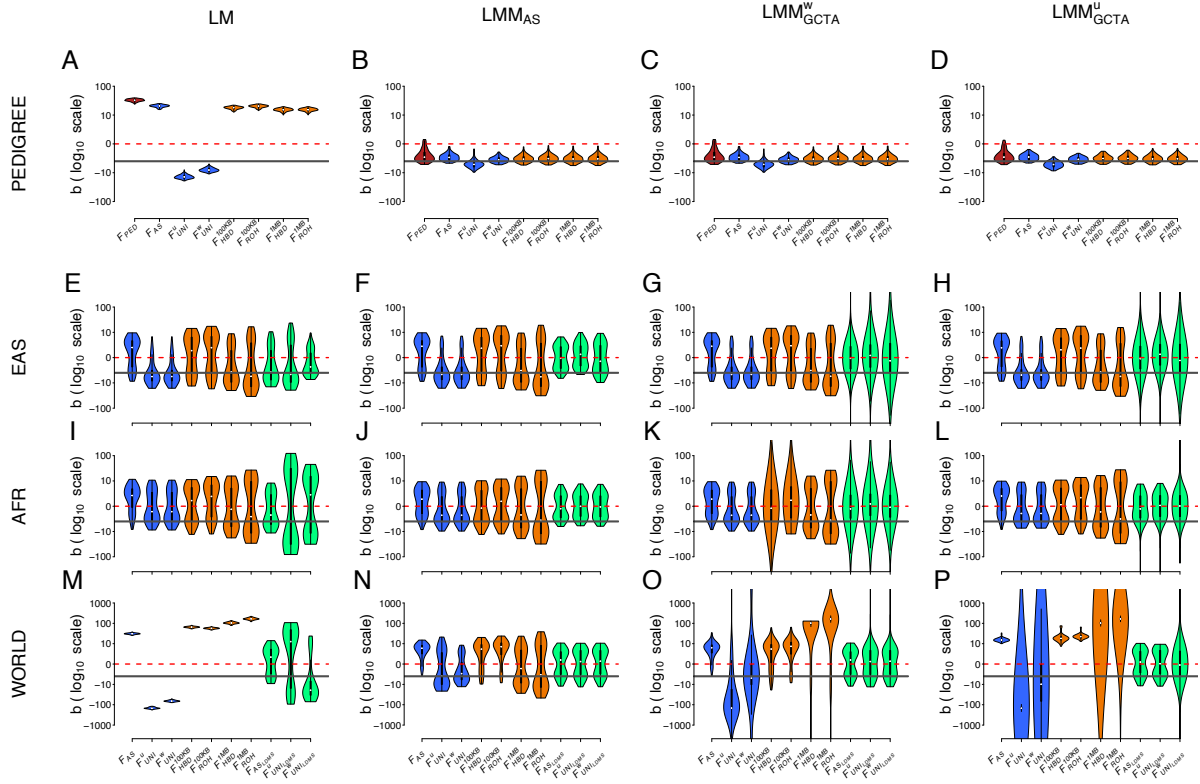


Figure S17: Comparison of the estimation of inbreeding depression strength (b) among different F estimates and models in four different populations with the ADD & DOM & DEMA scenario: effect sizes and dominance coefficients assigned proportional to the MAF of causal markers and presence of DEMA. Each column represents a regression model. The first column depicts the simple linear regression (panel A, E, I and M), the second column the linear mixed model with allele sharing GRM matrix as random factor (panel B, F, J and N), the third column the linear mixed model with the $GCTA^w$ relatedness matrix as random factor (panel C, G, K and O) and finally the fourth column represents the linear mixed model with $GCTA^u$ relatedness matrix as random factor (panel D, H, L and P). The first row depicts the complete simulated population (11,924 individuals): PEDIGREE in panels A, B, C and D. Inbreeding estimates compared in these panels (A – D) are F_{PED} , F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$ and $F_{ROH.1MB}$. The last three rows are the populations from the 1,000 Genomes Project: EAS in panels E, F, G and H, AFR in panels I, J, K and L and WORLD in panels M, N, O and P. Inbreeding estimates compared in these panels (E – P) are F_{AS} , F_{UNI}^u , F_{UNI}^w , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$, $F_{ROH.1MB}$, $F_{AS.LDMS}$, $F_{UNI.LDMS}^u$ and finally $F_{UNI.LDMS}^w$. Violin plots represent the distribution of the inbreeding depression strength estimates (b) among the 100 replicates. The solid dark grey line is the true strength of ID ($b = -3$). The dashed red line represents the absence of ID ($b = 0$), meaning that we failed to detect ID in any replicate above this line. Note that all panels (A – P) are in \log_{10} scale. Also note that linear mixed models did not converge for some replicates (yielding estimated b values above 1,000 or below -1,000, not shown if outside the graph limits). Percentages of replicates which did not converge: panel G (EAS, $GCTA^w$): 4% for $F_{AS.LDMS}$, 7% for $F_{UNI.LDMS}^u$ and 13% for $F_{UNI.LDMS}^w$; panel H (EAS, $GCTA^u$): 6% for $F_{AS.LDMS}$, 6% for $F_{UNI.LDMS}^u$ and 8% for $F_{UNI.LDMS}^w$; panel K (AFR, $GCTA^w$): 3% for $F_{HBD.100KB}$, 1% for $F_{ROH.100KB}$, 12% for $F_{AS.LDMS}$, 15% for $F_{UNI.LDMS}^u$ and 12% for $F_{UNI.LDMS}^w$; panel L (AFR, $GCTA^u$): 1% for $F_{UNI.LDMS}^u$, 1% for $F_{UNI.LDMS}^w$ and 1% for $F_{UNI.LDMS}^w$; panel O (WORLD, $GCTA^w$): 9% F_{UNI}^u , 21% F_{UNI}^w , 3% for $F_{HBD.1MB}$, 6% for $F_{ROH.1MB}$, 1% for $F_{UNI.LDMS}^u$ and 1% for $F_{UNI.LDMS}^w$; panel P (WORLD, $GCTA^u$) 14% F_{UNI}^u , 24% F_{UNI}^w , 13% for $F_{ROH.1MB}$, 6% for $F_{HBD.1MB}$ and finally 3% for $F_{UNI.LDMS}^w$.

Figure S17 shows the results of ID strength estimation for the ADD & DOM & DEMA scenario (when both the additive effects and dominance coefficients are inversely

proportional to MAF and there is presence of DEMA). Corresponding RMSE values can be found in tables S3 – S6. We find the same results as figure S16 with the LM in the complete PEDIGREE population: both F_{UNI} are overestimating the strength of ID while all the other F are largely underestimating it (panel A), suggesting that this pattern is due to the combination of the dominance coefficients being proportional to MAF and the presence of DEMA. In this scenario however, the three mixed models allow an efficient estimation of b with all the F (panels B – D, tables S3 – S6). There is no difference between the four models for both homogeneous populations (panels E – L, tables S3 – S6). Concerning the WORLD population, the smallest RMSE values were achieved only with the LMM_{AS} model and more specifically with F_{UNI}^W (panels M – P, tables S3 – S6). The variance among b estimates however are still fairly large, suggesting that we lack power (i.e. the sample size is too small). Finally, both $GCTA$ matrices resulted in biased b estimates (panels O and P, tables S5 and S6).

To summarize, we show that ADD, DOM and DEMA largely increase the variance around b estimates (particularly when more than one parameter was used). F_{UNI}^u and all IBD segments-based F were especially sensitive to the additive effect sizes being proportional to MAF. All F resulted in less accurate estimation of b when the dominance coefficients were proportional to MAF. Perhaps, this is because our current model is not accounting for dominance, despite knowing there is dominance. A possible solution not explored here might be to include an additional random factor with the dominance genomic relationships matrix [52]. Finally, F_{AS} and to a lesser extent F_{UNI}^u were strongly influenced by DEMA. This might be strange as Yengo *et al.* [56] showed that F_{UNI}^u was robust to DEMA but we show in figure S18 that it might be because they filtered on $MAF < 0.05$ in their analyses.

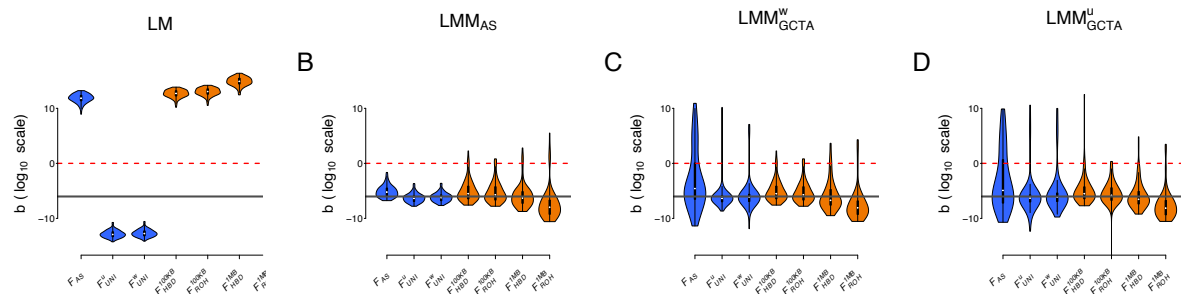


Figure S18: Comparison of the estimation of inbreeding depression strength (b) among different F estimates and models in the WORLD populations with the ADD & DOM & DEMA simulated scenario. SNPs have been filtered on $MAF > 0.05$: panel A shows the simple linear model; panel B shows the linear mixed model with the allele sharing relatedness matrix as random factor; panel C shows the linear mixed model with $GCTA^W$ relatedness matrix as random factor; panel D shows the linear mixed model with $GCTA^u$ relatedness

where IBD segments were inferred using `BCFTOOLS` and no size selection was performed, $F_{HBD.5MB}$ where HBD segments were called with `BCFTOOLS` and selected if larger than 5MB and $F_{ROH.5MB}$ where ROHs were called with `PLINK` and selected if larger than 5MB. The results of $F_{HBD.QUAL}$ are similar to those obtained with $F_{HBD.100KB}$ Concerning the F based on segments larger than 5MB, focusing on recent inbreeding did not improved the quantification of b .

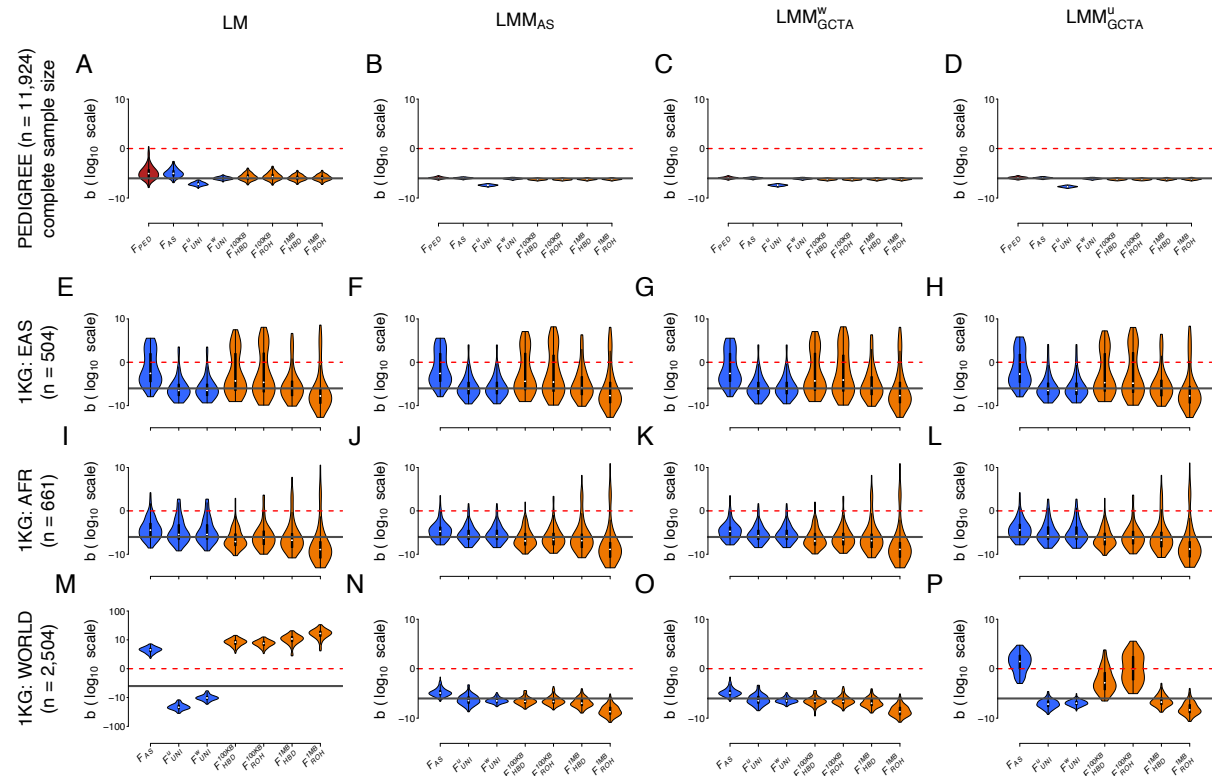


Figure S20: Comparison of the estimation of inbreeding depression strength (b) among different F estimates and models in four different populations with the ADD & DOM & DEMA scenario: effect sizes and dominance coefficients assigned proportional to the MAF of causal markers and presence of DEMA. Causal loci have been selected with intermediate frequencies: $MAF > 0.1$ Each column represents a regression model. The first column depicts the simple linear regression (panel A, E, I and M), the second column the linear mixed model with allele sharing GRM matrix as random factor (panel B, F, J and N), the third column the linear mixed model with the $GCTA^W$ relatedness matrix as random factor (panel C, G, K and O) and finally the fourth column represents the linear mixed model with $GCTA^U$ relatedness matrix as random factor (panel D, H, L and P). The first row depicts the complete simulated population (11,924 individuals): PEDIGREE in panels A, B, C and D. Inbreeding estimates compared in these panels (A – D) are F_{PED} , F_{AS} , F_{UNI}^U , F_{UNI}^W , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$ and $F_{ROH.1MB}$. The last three rows are the populations from the 1,000 Genomes Project: EAS in panels E, F, G and H, AFR in panels I, J, K and L and WORLD in panels M, N, O and P. Inbreeding estimates compared in these panels (E – P) are F_{AS} , F_{UNI}^U , F_{UNI}^W , $F_{HBD.100KB}$, $F_{ROH.100KB}$, $F_{HBD.1MB}$, and $F_{ROH.1MB}$. Violin plots represent the distribution of the inbreeding depression strength estimates (b) among the 100 replicates. The solid dark grey line is the true strength of ID ($b = -3$). The dashed red line represents the absence of ID ($b = 0$), meaning that we failed to detect ID in any replicate above this line. Note that all panels (A – P) are in \log_{10} scale. Also note that all linear models converged for all replicates.

As mentioned in the introduction, Alemu *et al.* (2021) [1] and Caballero *et al.* (2020) [7] showed that the best F actually depends on the history of the population. Indeed, they showed that F_{ROH} and F_{HBD} and to a lesser extent F_{HOM} were more efficient at quantifying homozygosity at loci with common alleles. On the contrary, F_{UNI}^u was better at quantifying homozygosity at rare alleles. The authors propose that in populations with low effective sizes, the strength of selection is diminished which can lead to deleterious alleles reaching intermediate frequencies because of drift. This means that both F_{ROH} and F_{HBD} will perform better in such populations. On the contrary, in populations with large effective size, selection maintains deleterious alleles a low frequencies which explains why Yengo *et al.* (2017) found that F_{UNI} was the best F with the large UK biobank dataset. Consequently, we simulated an additional scenario where intermediate frequencies causal loci were selected on $\text{MAF} > 0.1$ (figure S20). Firstly, this greatly reduced the variance around b estimates for all models and F . This suggests that it is much easier to detect inbreeding depression when the causal loci have higher frequencies. Secondly and as expected, it improved all IBD-segments based F (both F_{HBD} and both F_{ROH}) in all populations except $F_{\text{ROH}.100\text{KB}}$ and $F_{\text{HBD}.100\text{KB}}$ in the EAS population. This is because the EAS population has a small effective population size, resulting in large numbers of small IBD segments. Since inbreeding depression is mostly caused by recent coalescence events, including these smaller segments added noise in the models, leading to biased b estimates. This is consistent with previous studies who showed that only including larger fragments resulted in better inbreeding depression estimates [7, 6]. Finally, and surprisingly, excluding rare causal loci did not worsen F_{UNI}^w estimation of b as we would have expected. This might be because we did not explicitly simulate small populations, but simply selected common causal loci.

House sparrow morphological traits analyses

All the traits but bill depth have a negative slope b for the effect of F_{UNI}^w (table S3), hinting at the presence of Inbreeding depression for these traits. However, only bill length shows a strong and significant effect of F_{UNI}^w on all analyses, with a slope b around -1. For tarsus length, the standard LM or the mixed model without GRM gave an estimated slope b of -0.4 and -0.506 respectively, both non-significantly different from 0. With LMM_{AS} , $b = -0.771$ and the associated probability (p-value) is 0.047, while with the full model, $b = -0.699$ and $P_F = 0.071$. Thus, adding the GRM_{AS} for this trait led to a steeper slope and smaller p-values. For wing length, we see the reverse pattern, including the GRM_{AS} led to

a shallower slope and larger, non-significant p-values. For bill depth, the only trait with a positive slope for the effect of inbreeding, adding the GRM_{AS} makes the slope shallower and the p-value larger, while for mass, adding the GRM_{AS} makes the slopes steeper and the p-value smaller. For all of these traits but bill length, analyses with or without the GRM_{AS} would lead to different results and interpretations. Obviously, with empirical data, we cannot say which model is correct as we don't know the true value of the slope, but while the models we used are not as sophisticated as in the original publication, our conclusions for the presence of ID in models with GRM_{AS} would have been different from those of Niskanen *et al.*: we find significant inbreeding depression for adult mass, bill length and possibly tarsus length.

Table S2: Analyses of adults' tarsus length, wing length, bill depth, bill length and mass from 1,786 adult sparrows. LM: simple linear model with Sex and F_{UNI}^w as explanatory variables. LMM_{AS}: linear mixed model with sex and F_{UNI}^w as fixed effects and GRM_{AS} as random effect. LMM: linear mixed model with sex and F_{UNI}^w as fixed effects and island and year nested in island as random effects. LMM_{FULL}: linear mixed model with sex and F_{UNI}^w as fixed effects and island, year nested in island and GRM_{AS} as random effects. V_I : variance component of island effect; $V_{Y:I}$: variance component for year nested in island; V_A : additive variance; V_E : residual variance; P_F : p-value for the slope b of F_{UNI}^w to be 0.

Tarsus	Int.	Sex	F	V_{ISL}	$V_{Y:ISL}$	V_A	V_E	P_F
LM	19.443	0.074	-0.400				0.652	0.286
LMM _{AS}	19.503	0.081	-0.771			0.255	0.393	0.048
LMM	19.459	0.073	-0.506	0.000	0.045		0.607	0.179
LMM _{FULL}	19.513	0.080	-0.699	0.000	0.028	0.245	0.373	0.071
Wing	Int.	Sex	F	V_{ISL}	$V_{Y:ISL}$	V_A	V_E	P_F
LM	78.548	2.610	-1.695				3.073	0.037
LMM _{AS}	76.408	2.605	-1.437			1.645	1.304	0.068
LMM	78.392	2.591	-1.405	0.274	0.069		2.825	0.087
LMM _{FULL}	76.331	2.614	-1.280	0.064	0.068	1.619	1.245	0.103
Bill depth	Int.	Sex	F	V_{ISL}	$V_{Y:ISL}$	V_A	V_E	P_F
LM	8.125	0.035	0.265				0.076	0.039
LMM _{AS}	8.084	0.032	0.216			0.035	0.044	0.106
LMM	8.118	0.035	0.238	0.001	0.006		0.071	0.068
LMM _{FULL}	8.073	0.034	0.227	0.000	0.005	0.036	0.039	0.084
Bill length	Int.	Sex	F	V_{ISL}	$V_{Y:ISL}$	V_A	V_E	P_F
LM	13.75	-0.010	-1.172				0.285	0.000
LMM _{AS}	13.59	-0.013	-1.066			0.140	0.135	0.000
LMM	13.68	-0.016	-0.953	0.017	0.020		0.254	0.000
LMM _{FULL}	13.60	-0.012	-1.002	0.007	0.009	0.124	0.135	0.000
Mass	Int.	Sex	F	V_{ISL}	$V_{Y:ISL}$	V_A	V_E	P_F
LM	32.931	-1.387	-2.390				4.591	0.016

LMM _{AS}	34.324	-1.414	-2.856			1.560	3.019	0.007
LMM	32.939	-1.377	-1.983	0.149	0.269		4.269	0.050
LMM _{FULL}	34.341	-1.395	-2.851	0.099	0.172	1.445	2.915	0.006

SI Dataset can be found on [GitHub](#).

Dataset S1 (TablesS3-S6_ID_EST_RMSE_SUPP_TABLES.xlsx) contains the RMSE for all scenarios and all populations per inbreeding coefficient. RMSE are across the 100 simulation replicates.

Dataset S2 (TablesS7-S9_ID_EST_RepConvergence_SUPP_TABLES.xlsx) contains the percentage of replicates which did not converge in the three LMM for all scenarios and all populations per inbreeding coefficient. Percentages are across the 100 simulation replicates.

References

- [1] S.W. Alemu, N.K. Kadri, C. Harland, P. Faux, C. Charlier, A. Caballero, and T. Druet. An evaluation of inbreeding measures using a whole-genome sequenced cattle pedigree. *Heredity*, 126(3):410–423, 2021.
- [2] A.H. Bittles and M.L. Black. Consanguinity, human evolution, and complex diseases. *Proceedings of the National Academy of Sciences*, 107(suppl 1):1779–1786, 2010.
- [3] A.H. Bittles. *Consanguinity in Context*. Cambridge University Press, 1 edition, 2012.
- [4] B.L. Browning and S.R. Browning. Detecting Identity by Descent and Estimating Genotype Error Rates in Sequence Data. *The American Journal of Human Genetics*, 93(5):840–851, 2013.
- [5] C. Béréanos, P.A. Ellis, J.G. Pilkington, and J.M. Pemberton. Genomic analysis reveals depression due to both individual and maternal inbreeding in a free-living mammal population. *Molecular Ecology*, 25(13):3152–3168, 2016.
- [6] A. Caballero, A. Fernández, B. Villanueva, and M.A. Toro. A comparison of marker-based estimators of inbreeding and inbreeding depression. *Genetics Selection Evolution*, 54(1):82, 2022.
- [7] A. Caballero, B. Villanueva, and T. Druet. On the estimation of inbreeding depression using different measures of inbreeding from molecular markers. *Evolutionary Applications*, 14(2):416–428, 2021.
- [8] F.C. Ceballos, S. Hazelhurst, D.W. Clark, G. Agongo, G. Asiki, P.R. Boua, F.X. Gómez-Olivé, F. Mashinya, S. Norris, J.F. Wilson, and M. Ramsay. Autozygosity influences cardiometabolic disease-associated traits in the AWI-Gen sub-Saharan African study. *Nature communications*, 11(1):5754, 2020.
- [9] F.C. Ceballos, S. Hazelhurst, and M. Ramsay. Assessing runs of Homozygosity: a comparison of SNP Array and whole genome sequence low coverage data. *BMC Genomics*, 19(1):106, 2018.
- [10] F.C. Ceballos, P.K. Joshi, D.W. Clark, M. Ramsay, and J.F. Wilson. Runs of homozygosity: windows into population history and trait architecture. *Nature Reviews Genetics*, 19(4):220–234, 2018.
- [11] C.C. Chang, C.C. Chow, L.C. Tellier, S. Vattikuti, S.M. Purcell, and J.J. Lee. Second-generation *PLINK*: rising to the challenge of larger and richer datasets. *GigaScience*, 4(1):7, 2015.
- [12] D. Charlesworth and J.H. Willis. The genetics of inbreeding depression. *Nature Reviews Genetics*, 10(11):783–796, 2009.
- [13] A.W. Diamond. *Studies of Mascarene Island Birds*. Cambridge University Press, 1 edition, 1987.
- [14] H.P. Doekes, P. Bijma, and J. J. Windig. How Depressing Is Inbreeding? A Meta-Analysis of 30 Years of Research on the Effects of Inbreeding in Livestock. *Genes*, 12(6):926, 2021

- [15]T. Druet and M. Gautier. A model-based approach to characterize individual inbreeding at both global and local genomic scales. *Molecular Ecology*, 26(20):5820–5841, 2017.
- [16]R. Frankham. Conservation genetics. *Annual Review of Genetics*, 29(1):305–327, 1995.
- [17]S. Gazal, M. Sahbatou, M-C. Babron, E. Génin, and A-L. Leutenegger. High level of inbreeding in final phase of 1000 Genomes Project. *Scientific Reports*, 5(1):17453, 2015.
- [18]J. Goudet, T. Kay and, B.S. Weir. How to estimate kinship. *Molecular Ecology*, 27(20):4121–4135, 2018.
- [19]P.W. Hedrick. What is the evidence for heterozygote advantage selection? *Trends in Ecology & Evolution*, 27(12):698–704, 2012.
- [20]J. Huisman, L.E.B. Kruuk, P.A. Ellis, T. Clutton-Brock, and J.M. Pemberton. Inbreeding depression across the lifespan in a wild mammal population. *Proceedings of the National Academy of Sciences*, 113(13):3585–3590, 2016.
- [21]M. Kardos, G. Luikart, and F.W. Allendorf. Measuring individual inbreeding in the age of genomics: marker-based measures are better than pedigrees. *Heredity*, 115(1):63–72, 2015.
- [22]M. Kardos, P. Nietlisbach, and P.W. Hedrick. How should we compare different genomic estimates of the strength of inbreeding depression? *Proceedings of the National Academy of Sciences*, 115(11), 2018.
- [23]M. Kardos, Y. Zhang, K. M. Parsons, A. Yunga, H. Kang, X. Xu, X. Liu, C.O. Matkin, P. Zhang, E.J. Ward, M.B. Hanson, C. Emmons, M.J. Ford, G. Fan, and S. Li. Inbreeding depression explains killer whale population dynamics. *Nature Ecology & Evolution*, 7(5):675–686, 2023.
- [24]M. Kardos, M. Åkesson, T. Fountain, Ø. Flagstad, O. Liberg, P. Olason, H. Sand, P. Wabakken, C. Wikenros, and H. Ellegren. Genomic consequences of intensive inbreeding in an isolated wolf population. *Nature Ecology & Evolution*, 2(1):124–131, 2017.
- [25]L. Keller, and D.G. Waller. Inbreeding effects in wild populations. *Trends in Ecology & Evolution*, 17(5):230–241, 2002.
- [26]L. Keller, J.M Reid, and P. Arcese. Testing evolutionary models of senescence in a natural population: age and inbreeding effects on fitness components in song sparrows. *Proceedings of the Royal Society B: Biological Sciences*, 275(1635):597–604, 2008.
- [27]M.C. Keller, P.M. Visscher, and M.E. Goddard. Quantification of Inbreeding Due to Distant Ancestors and Its Detection Using Dense Single Nucleotide Polymorphism Data. *Genetics*, 189(1):237–249, 2011.
- [28]E-S. Kim, J.B. Cole, H. Huson, G.R. Wiggans, P. C. van Tassell, B.A. Crooker, G. Liu, Y. Da, and T.S. Sonstegard. Effect of Artificial Selection on Runs of Homozygosity in U.S. Holstein Cattle. *PLoS ONE*, 8(11):e80813, 2013.
- [29]D.G. Kleiman, B.B. Beck, J.M. Dietz, L.A. Dietz, J.D. Ballou, and A.F. Coimbra-Filho. Conservation Program for the Golden Lion Tamarin: Captive Research and Management, Ecological Studies, Educational Strategies, and Reintroduction. *Primates. Proceedings in Life Sciences*. 959–979, 1986.
- [30]E. Lavanchy and J. Goudet. Effect of reduced genomic representation on using runs of homozygosity for inbreeding characterization. *Molecular Ecology Resources*, 23(4):787–802, 2023.
- [31]A-L. Leutenegger, B. Prum, E. Génin, C. Verny, A. Lemainque, F. Clerget-Darpoux, and E.A. Thompson. Estimation of the Inbreeding Coefficient through Use of Genomic Data. *The American Journal of Human Genetics*, 73(3):516–523, 2003.
- [32]G. Malécot *Les mathématiques de l'hérédité*. Masson & Cie Paris, Paris, 1948.
- [33]K. Martikainen, A. Sironen, and P. Uimari. Estimation of intrachromosomal inbreeding depression on female fertility using runs of homozygosity in Finnish Ayrshire cattle. *Journal of Dairy Science*, 101(12):11097–11107, 2018.
- [34]K. Martikainen, A.M. Tyrisevä, K. Matilainen, J.P. Pösö, and P. Uimari. Estimation of inbreeding depression on female fertility in the Finnish Ayrshire population. *Journal of Animal Breeding and Genetics*, 134(5):383–392, 2017.
- [35]S. Mastrangelo, M. Tolone, R. Di Gerlando, L. Fontanesi, M.T. Sardina, and B. Portolano. Genomic inbreeding estimation in small populations: evaluation of runs of homozygosity in three local dairy cattle breeds. *Animal*, 10(5):746–754, 2016.

- [36]R. McQuillan, A-L. Leutenegger, R. Abdel-Rahman, C.S. Franklin, M. Pericic, L. Barac-Lauc, N. Smolej-Narancic, B. Janicijevic, O. Polasek, A. Tenesa, A.K. MacLeod, S.M. Farrington, P. Rudan, C. Hayward, V. Vitart, I. Rudan, S.H. Wild, M.G. Dunlop, A.F. Wright, H. Campbell, and J.F. Wilson. Runs of Homozygosity in European Populations. *The American Journal of Human Genetics*, 83(3):359–372, 2008.
- [37]R. Meyermans, W. Gorssen, N. Buys, and S. Janssens. How to study runs of homozygosity using *PLINK*? A guide for analyzing medium density SNP data in livestock and pet species. *BMC Genomics*, 21(1):94, 2020.
- [38]V. Narasimhan, P. Danecek, A. Scally, Y. Xue, C. Tyler-Smith, and R. Durbin. *BCFtools/RoH*: a hidden Markov model approach for detecting autozygosity from next-generation sequencing data. *Bioinformatics*, 32(11):1749–1751, 2016.
- [39]P. Nietlisbach, S. Muff, J.M. Reid, M.C. Whitlock, and L. F. Keller. Nonequivalent lethal equivalents: Models and inbreeding metrics for unbiased estimation of inbreeding load. *Evolutionary Applications*, 12(2):266–279, 2019.
- [40]P.A. Oltenacu and B. Algers. Selection for increased production and the welfare of dairy cows: are new breeding goals needed? *Ambio*, 34(4- 5):311–315, 2005.
- [41]E. Peripolli, N. Bonvino Stafuzza, D. Prado Munari, A.L. Ferreira Lima, R. Irgang, M. Antonio Machado, J.C. Do Carmo Panetto, R. Vieira Ventura, F. Baldi, and M.V. Gualberto Barbosa Da Silva. Assessment of runs of homozygosity islands and estimates of genomic inbreeding in Gyr (*Bos indicus*) dairy cattle. *BMC Genomics*, 19(1):34, 2018.
- [42]E. Postma, L. Martini, and P. Martini. Inbred women in a small and isolated Swiss village have fewer children. *Journal of Evolutionary Biology*, 23(7):1468–1474, 2010.
- [43]J.K. Pritchard. Are Rare Variants Responsible for Susceptibility to Complex Diseases? *The American Journal of Human Genetics*, 69(1):124– 137, 2001.
- [44]J.E. Pryce, M. Haile-Mariam, M.E. Goddard, and B.J. Hayes. Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle. *Genetics Selection Evolution*, 46(1):71, 2014.
- [45]S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M.A.R. Ferreira, D. Bender, J. Maller, P. Sklar, P.I.W. de Bakker, M.J. Daly, and P.C. Sham. *PLINK*: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, 81(3):559–575, 2007.
- [46]K. Ralls and J.D. Ballou. Genetic Status and Management of California Condors. *The Condor*, 106(2):215–228, 2004.
- [47]J.A. Robinson, C. Brown, B.Y. Kim, K.E. Lohmueller, and R.K. Wayne. Purging of Strongly Deleterious Mutations Explains Long-Term Persistence and Absence of Inbreeding Depression in Island Foxes. *Current Biology*, 28(21):3487–3494.e4, 2018.
- [48]R.M. Santymire, T.M. Livieri, H. Branvold-Faber, and P.E. Marinari. The Black-Footed Ferret: On the Brink of Recovery? In William V. Holt, Janine L. Brown, and Pierre Comizzoli, editors, *Reproductive Sciences in Animal Conservation*, volume 753, pages 119–134, 2014.
- [49]M.A. Stoffel, S.E. Johnston, J.G. Pilkington, and J.M. Pemberton. Genetic architecture and lifetime dynamics of inbreeding depression in a wild mammal. *Nature Communications*, 12(1):2972, 2021.
- [50]The International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. *Nature*, 449(7164):851–861, 2007.
- [51]K.J.H. Verweij, A. Abdellaoui, J. Veijola, S. Sebert, M. Koiranen, M.C. Keller, M-R. Järvelin, and B.P. Zietsch. The Association of Genotype-Based Inbreeding Coefficient with a Range of Physical and Psychological Human Traits. *PLoS ONE*, 9(7):e103102, 2014.
- [52]Z.G. Vitezica, L. Varona, and A. Legarra. On the Additive and Dominant Variance and Covariance of Individuals Within the Genomic Selection Scope. *Genetics*, 195(4):1223–1230, 2013.
- [53]S. Wright. Coefficients of Inbreeding and Relationship. *The American Naturalist*, 56(645):330–338, 1922.
- [54]J. Yang, S.H. Lee, M.E. Goddard, and P.M. Visscher. *GCTA*: A Tool for Genome-wide Complex Trait Analysis. *The American Journal of Human Genetics*, 88(1):76–82, 2011.

- [55]L. Yengo, J. Yang, M.C. Keller, M. E. Goddard, N.R. Wray, and P.M. Visscher. Genomic partitioning of inbreeding depression in humans. *The American Journal of Human Genetics*, 108(8):1488– 1501, 2021.
- [56]L. Yengo, Z. Zhu, N.R. Wray, B.S. Weir, J.Yang, M.R. Robinson, and P.M. Visscher. Detection and quantification of inbreeding depression for complex traits from SNP data. *Proceedings of the National Academy of Sciences*, 114(32):8602–8607, 2017.
- [57]C. Zhang, P. Wang, D.Tang, Z. Yang, F. Lu, J. Qi, N.R. Tawari, Y. Shang, C. Li, and S. Huang. The genetic basis of inbreeding depression in potato. *Nature Genetics*, 51(3):374–378, 2019.
- [58]Q.S. Zhang, J. Goudet, and B.S. Weir. Rank-invariant estimation of inbreeding coefficients. *Heredity*, 128(1):1–10, 2022
- [59]B.S. Weir, and J. Goudet. A Unified Characterization of Population Structure and Relatedness. *Genetics*, 206(4):2085–2103, 2017.
- [60]A.K. Niskanen, A.M. Billing, H. Holand, I.J. Hagen, Y.G. Araya-Ajoy, A. Husby, B. Rønning, A.M. Myhre, P.S. Ranke, T. Kvalnes, H. Pärn, T.H. Ringsby, S. Lien, B-E. Sæther, S. Muff, and H. Jensen. Consistent scaling of inbreeding depression in space and time in a house sparrow metapopulation. *Proceedings of the National Academy of Sciences*, 117(25):14584–14592, 2020.

Supplementary Material Chapter III: Too big to purge: no evidence of purging in the islands of the European barn owl (*Tyto alba*)

Authors: Eléonore Lavanchy, Tristan Cumer, Alexandros Topaloudis, Anne-Lyse Ducrest, Céline Simon, Alexandre Roulin and Jérôme Goudet

Supplementary Material and Methods

Trimming relatedness

The 502 individuals were trimmed for allele-sharing based relatedness > 0.05 with a custom R script and the `hierfstat` R package. 187 individuals were left after relatedness trimming.

Estimating additional inbreeding coefficients

F_{PED} was estimated as twice the diagonal of the kinship matrix minus 1. The kinship matrix was estimated with the `kinship` function from the `kinship2` R package.

Comparing minor and derived allele identification

As mentioned in the main text, we compare the fraction of sites for which the minor allele (we identified by bootstrapping individuals) corresponds to the derived allele in a set of 1,373,932 neutral sites. We show that the mean fraction of correct assignment is 87% but that this fraction decreases linearly with site's MAF (figure S1).

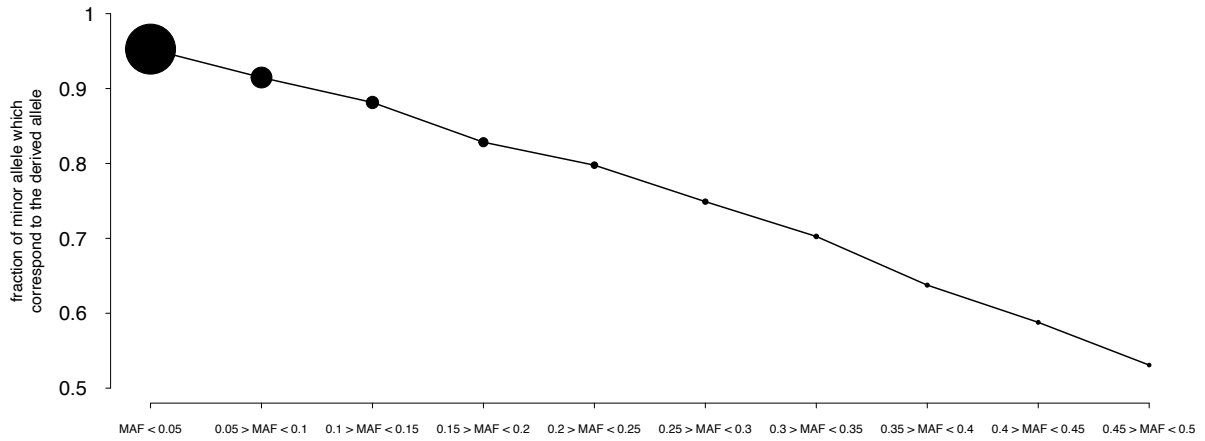


Figure S1: Fraction of minor allele which correspond to the derived allele. Minor allele was estimated globally via 1,000 bootstraps by sampling only unrelated individuals. The size of the dots corresponds to the number of alleles in each MAF category.

Supplementary Results

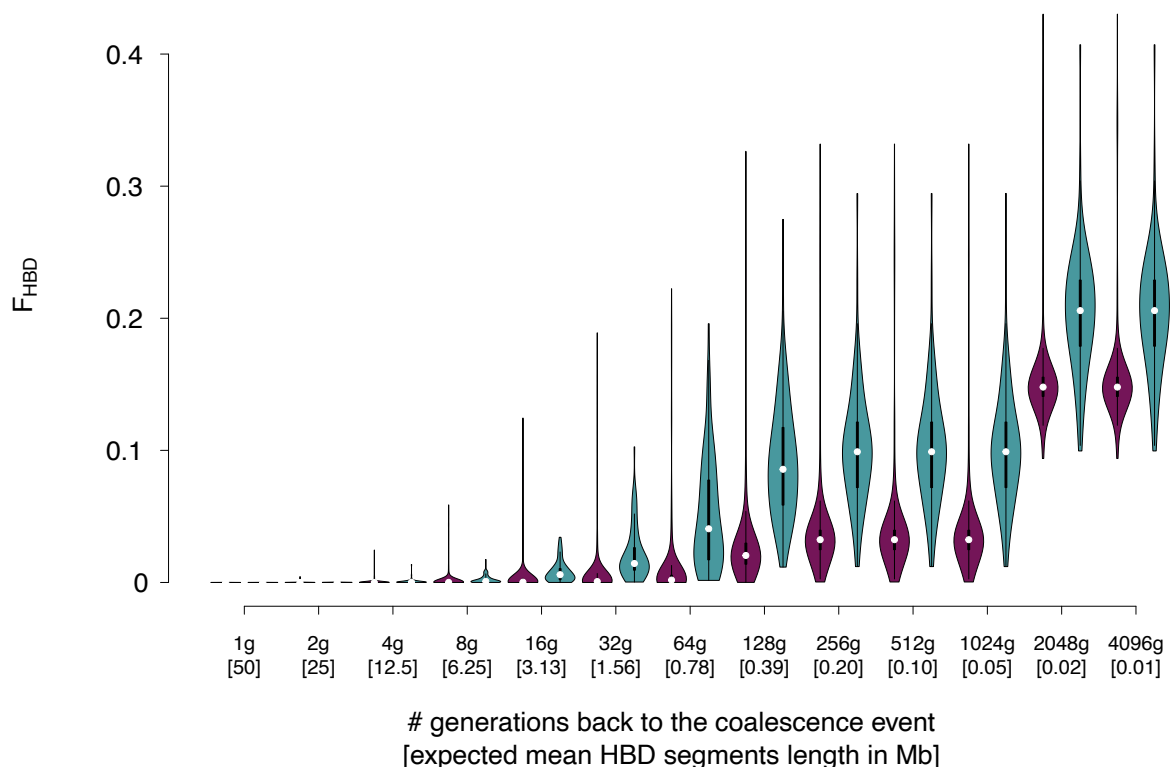


Figure S2: Cumulative F_{HBD} (estimated as the average (among variants) probability that a variant belongs to an HBD segment coalescing less than 512 generations ago) distributions among the different HBD classes used in $RZ_{OO}R_{OH}$. Continental populations are colored in purple and island populations in blue.

This plot shows the cumulative F_{HBD} across the different HBD classes used in the $RZ_{OO}R_{OH}$ model. We show that HBD segments coalescing less than 8 generations ago are

very rare and mostly present in the truly inbred individuals mentioned above. Difference between continental and islands populations start to show from HBD segments coalescing around 16 generations ago. Islands F_{HBD} especially increase when coalescence events from 64 generations are included. In the main text analyses, we only consider SNPs or segments HBD when they are included in the first 10 classes (from 1 generation ago to 512 generations ago).

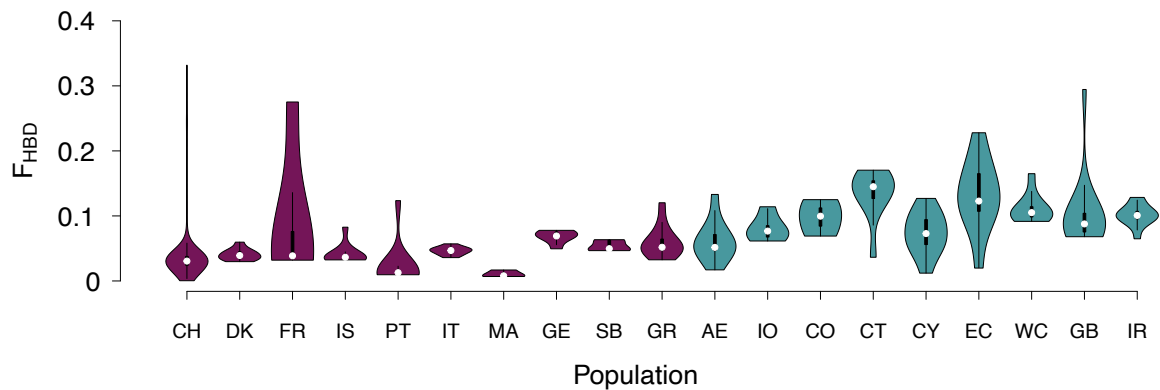


Figure S3: F_{HBD} (estimated as the average (among variants) probability that a variant belongs to an HBD segment coalescing less than 512 generations ago) distributions per population. Continental populations are colored in purple and island populations in blue.

On average, islands populations show higher F_{HBD} compared to continental populations. We detect few truly inbred individuals (resulting from mating between close relatives): 16 from CH, 1 from PT, one from FR, one from GB and one from EC. Among continental populations, individuals from GE, SB and GR shows highest F_{HBD} . Among island populations, individuals from AE and CY showed the lowest F_{HBD} .

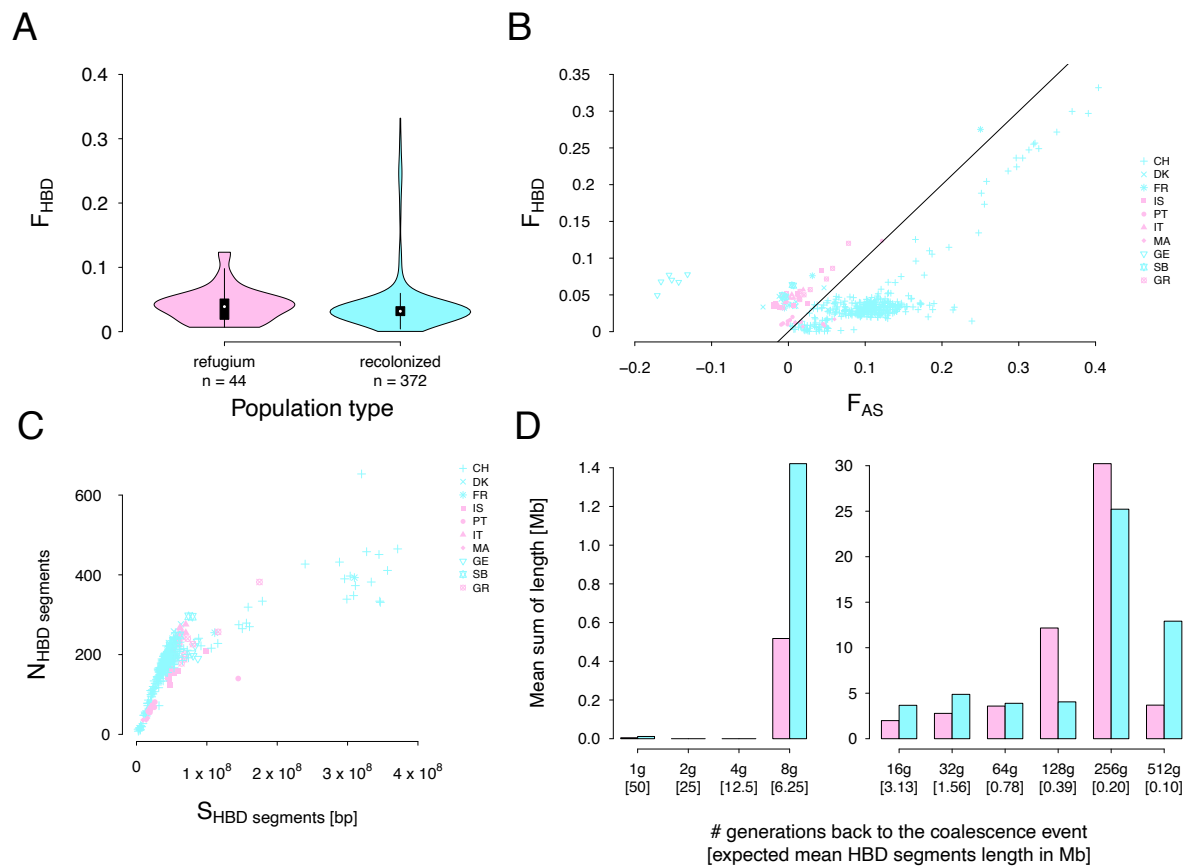


Figure S4: For all panels, refugia populations are shown in pink and recolonized populations in blue. **A:** F_{HBD} distributions from refugia and recolonized populations. F_{HBD} considers a marker as autozygous if the coalescence event is before 512 generations ago. **B:** scatter plot of F_{HBD} against F_{AS} . Each point represents one individual and its shape indicates which population it comes from. The black line is the identity line ($x = y$). **C:** number of HBD segments (N_{HBD}) as a function of the mean length of HBD segments (S_{HBD}) in base-pair. Each point represents one individual and its shape indicates which population it comes from. **D:** HBD segments distributions from refugia populations and recolonized populations. The y-axis represents the mean sum of length (among individuals) falling into the different categories of HBD segments (represented in the x-axis).

There is little (but significant) difference between continental refugia and recolonized populations F_{HBD} distribution (Wilcoxon rank-sum test: $W = 10'063$, p-value = 0.01275; effect size: 0.122, considered small). For instance, there is no strongly inbred individuals in refugia populations. In addition, both refugia and recolonized populations are equally close to the F_{HBD} and F_{AS} identity line (except the Swiss individuals which are largely below). Refugia populations tended to have smaller HBD fragments for the same fraction of genome within HBD segments compared to recolonized populations (except for one inbred Greek individual). Finally, the refugia populations were slightly enriched in small HBD segments (coalescing 128 and 256 generations ago). On the contrary recolonized populations were enriched in both large (coalescing 8 generations ago) and

very small (coalescing 512 generations ago) HBD segments but the enrichment in large HBD segments is solely due to the few inbred Swiss present in the recolonized populations.

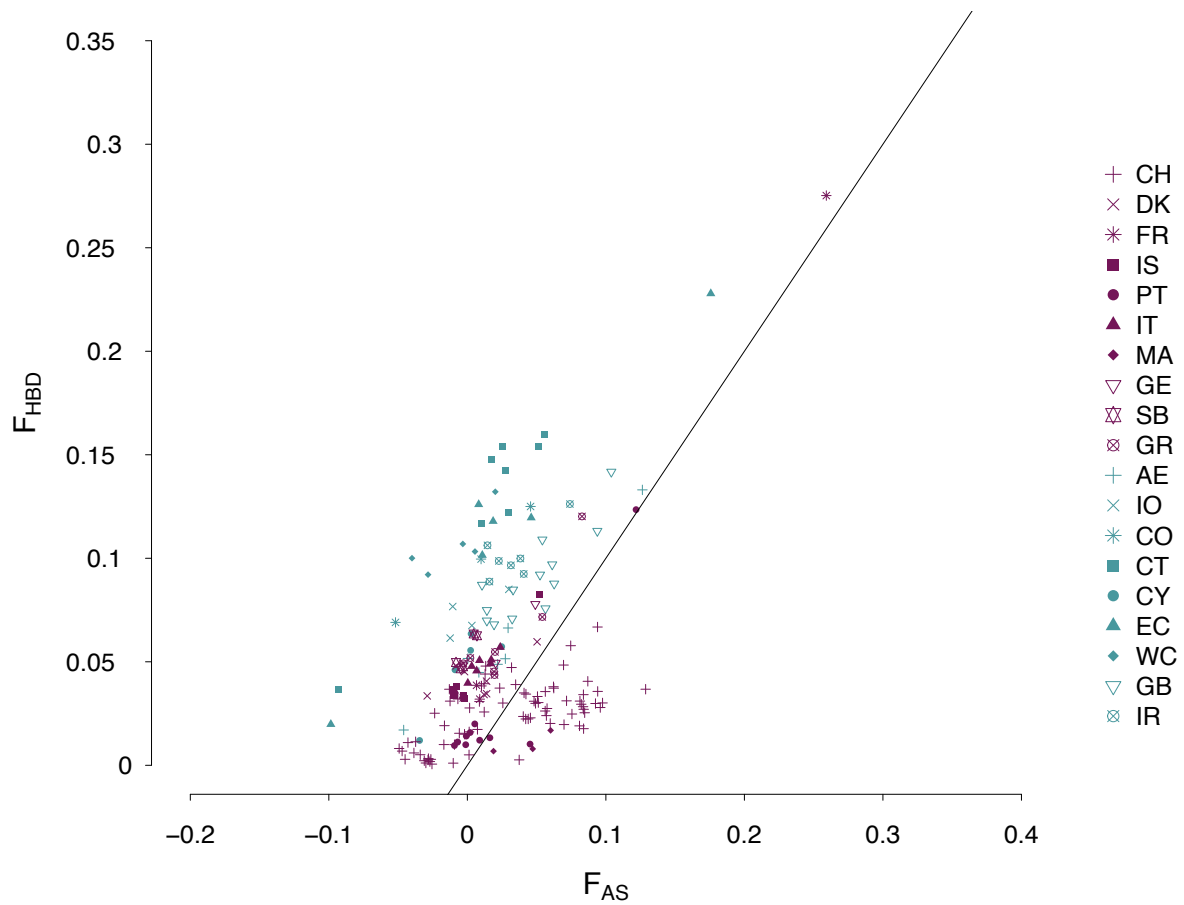


Figure S5: F_{HBD} according to F_{AS} in the set of 187 unrelated individuals. Each point represents one individual and its shapes indicates the population it comes from. The black line represents the identify line ($x = y$). Continental populations are colored in purple and island populations in blue.

This plot compares F_{HBD} and F_{AS} inbreeding coefficients with the unrelated set of individuals. Most individuals below the line have been removed or shifted towards the line, however few individuals are still below the line. We believe it is due to that fact that we did not filter on relatedness 0 (but rather 0.05). In addition, the Swiss sample we have is much larger compared to other populations and is known to contain families. We also know that MA samples are somehow related.

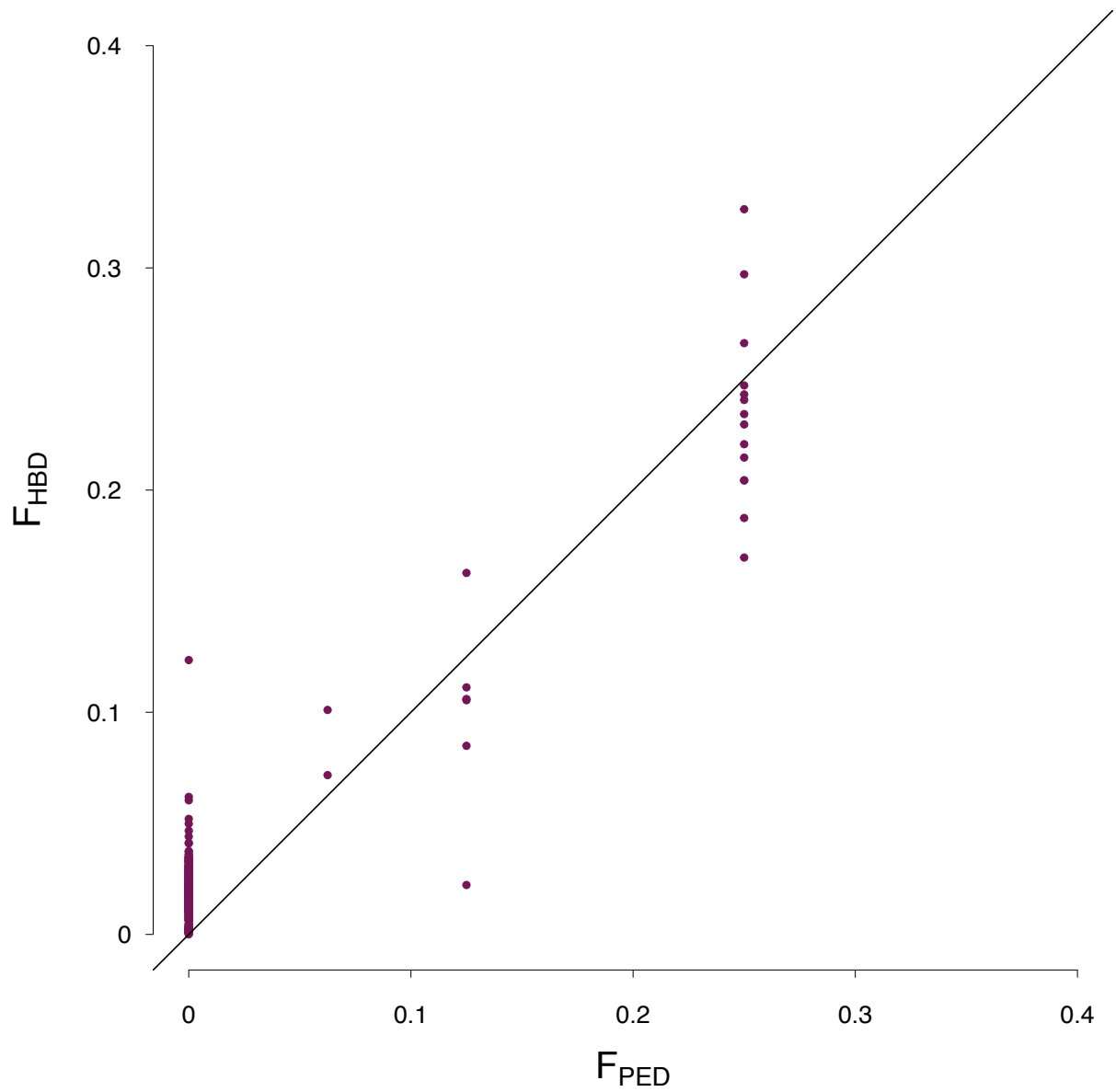


Figure S6: Comparison between F_{HBD} (estimated as the average (among variants) probability that a variant belongs to an HBD segment coalescing less than 512 generations ago) and F_{PED} in the Swiss population (CH).

This plot shows the comparison between the genetic observed F_{HBD} and the expected F_{PED} for the CH population. We show that there is a good correlation between both inbreeding coefficients (0.94).

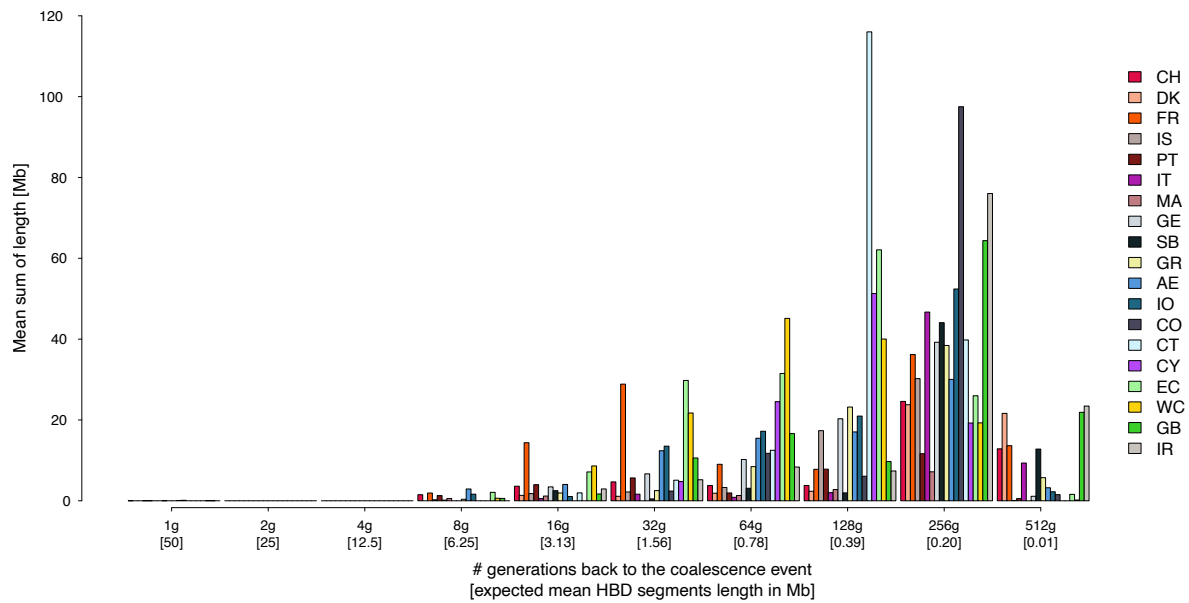


Figure S7: HBD segments distributions per population. Populations order is as follow: CH, DK, FR, IS, PT, IT, MA, GE, SB, GR, AE, IO, CO, CT, CY, EC, WC, GB, IR. The y axis represents the mean sum of length (among individuals) falling into the different categories of HBD segments (represented in the x axis).

We show HBD segments distributions among the different populations. In populations with small sample sizes, the peaks of mean sum of lengths are mostly driven by inbred individuals (namely in FR for HBD classes 5 and 6 and EC for HBD class 6 and 7). We observe that CT population has an especially high peak in the 8th HBD class coherent with its history of isolation and small effective population size.

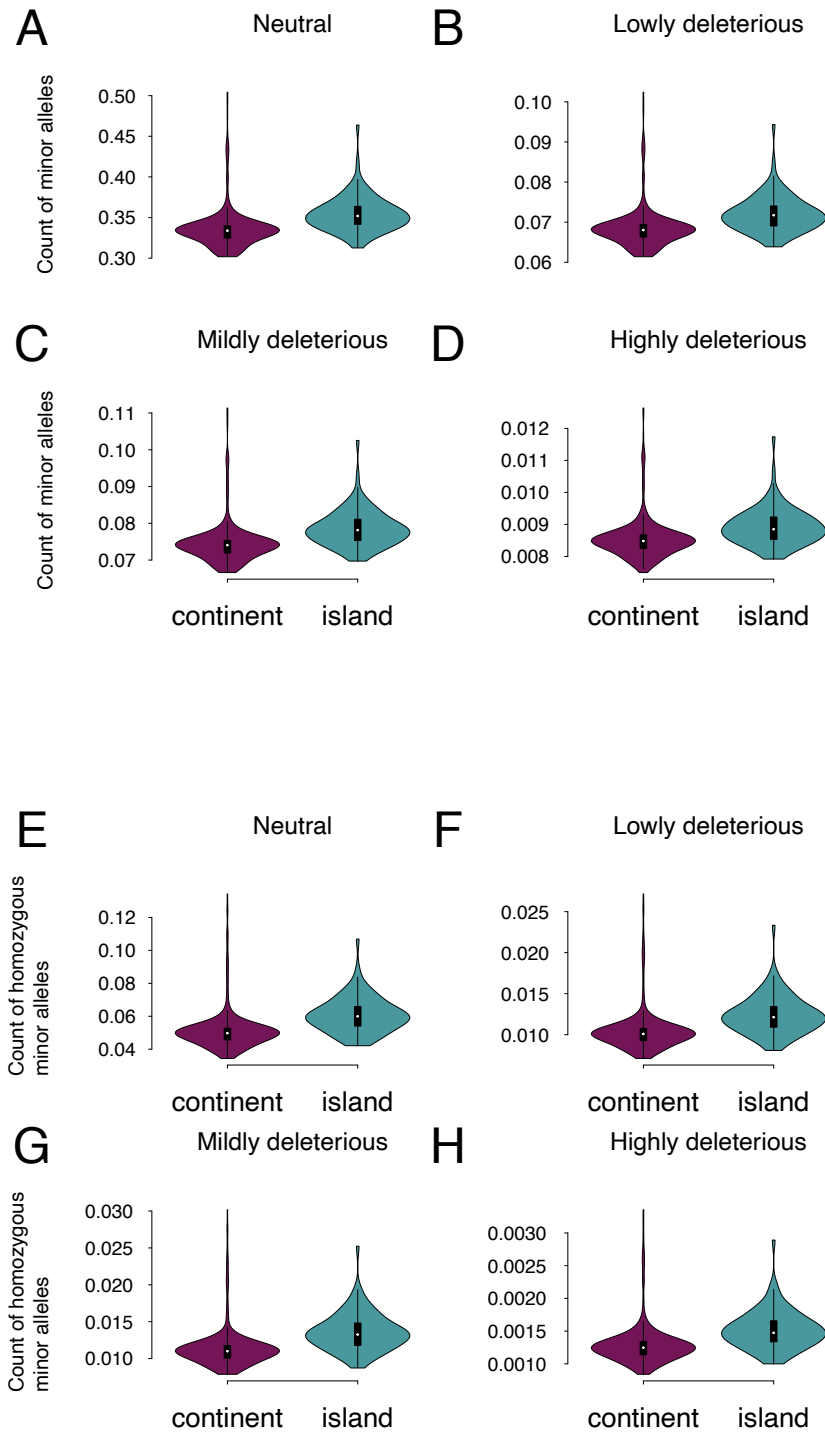


Figure S8: Distribution of minor alleles in continental versus islands populations. Minor alleles effects were classified with *SNPeff*. **A:** Count of neutral minor alleles divided by the individual number of polymorphic sites. **B:** Count of lowly deleterious minor alleles divided by the individual number of polymorphic sites. **C:** Count of moderately deleterious minor alleles divided by the individual number of polymorphic sites. **D:** Count of highly deleterious minor alleles divided by the individual number of polymorphic sites. **E:** Count of homozygous neutral minor alleles divided by the individual number of polymorphic sites. **F:** Count of homozygous lowly deleterious minor alleles divided by the individual number of polymorphic sites. **G:** Count of homozygous moderately deleterious minor alleles divided by the individual number of polymorphic

sites. **H**: Count of homozygous highly deleterious minor alleles divided by the individual number of polymorphic sites.

This plot shows the individual count of minor alleles divided by the individual number of polymorphic sites per variants category (A: Neutral; B: Lowly deleterious, C: Moderately deleterious; D: Highly deleterious; E: homozygous neutral; F: Homozygous lowly deleterious, G: Homozygous moderately deleterious; H: Homozygous highly deleterious) in continental populations versus island populations. Islands populations are significantly enriched in all types of minor alleles both for single copy and homozygous states (Neutral single copy: $W = 6904$, $p\text{-value} < 2.2e-16$ Lowly deleterious single copy: $W = W = 7739$, $p\text{-value} < 2.2e-16$; Mildly deleterious single copy: $W = 7410$, $p\text{-value} < 2.2e-16$; Highly deleterious single copy: $W = 9167$, $p\text{-value} = 1.07e-12$; Neutral homozygous: $W = 6752$, $p\text{-value} < 2.2e-16$; Lowly deleterious single copy: $W = 6748$, $p\text{-value} < 2.2e-16$; Mildly deleterious single copy: $W = 6613$, $p\text{-value} < 2.2e-16$; Highly deleterious single copy: $W = 7555$, $p\text{-value} < 2.2e-16$).

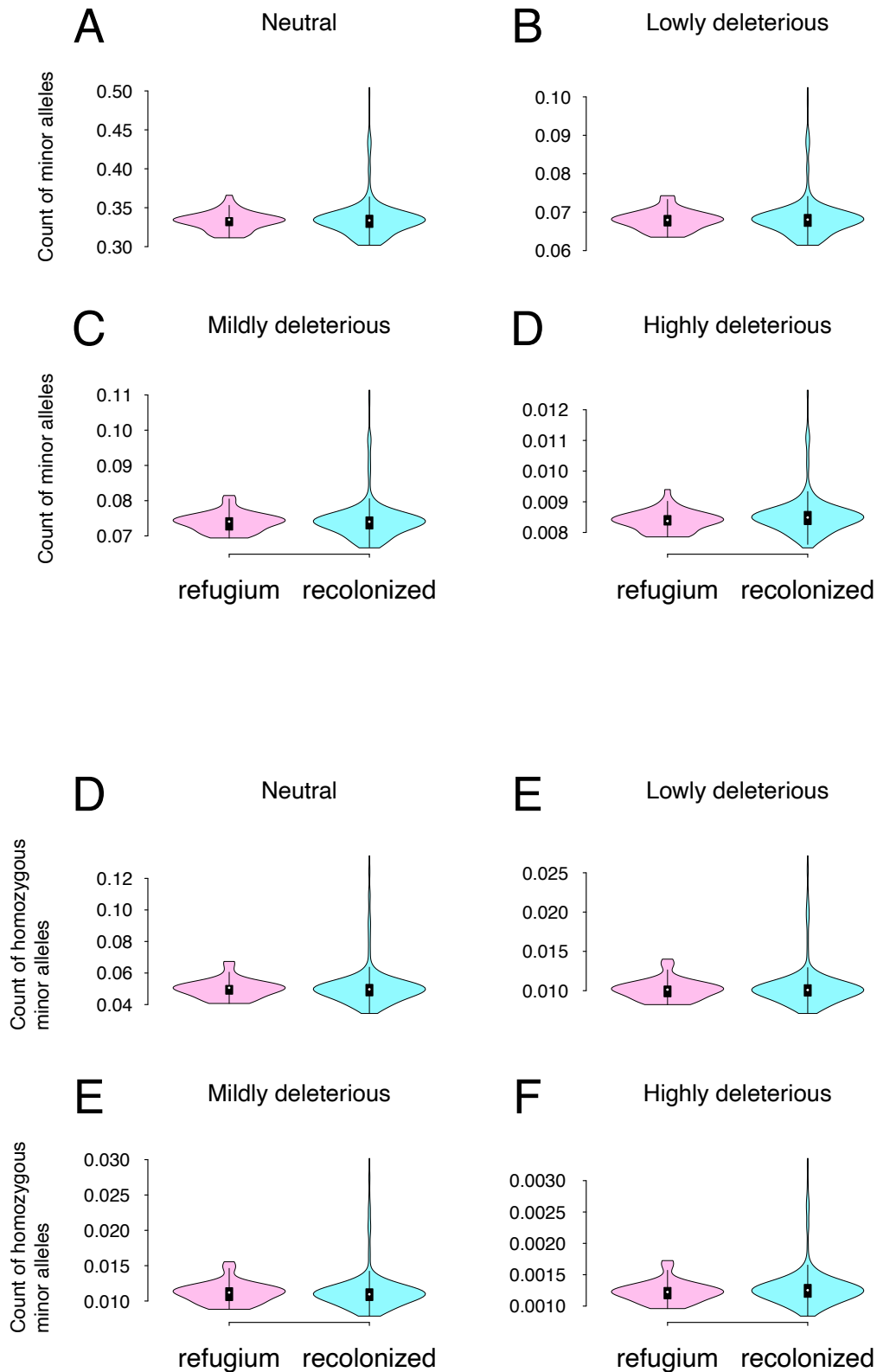


Figure S9: Distribution of minor alleles in refugium versus recolonized populations. Minor alleles effects were classified with *SNPeff*. **A:** Count of neutral minor alleles divided by the individual number of polymorphic sites. **B:** Count of lowly deleterious minor alleles divided by the individual number of polymorphic sites. **C:** Count of moderately deleterious minor alleles divided by the individual number of

polymorphic sites. **D**: Count of highly deleterious minor alleles divided by the individual number of polymorphic sites. **E**: Count of homozygous neutral minor alleles divided by the individual number of polymorphic sites. **F**: Count of homozygous lowly deleterious minor alleles divided by the individual number of polymorphic sites. **G**: Count of homozygous moderately deleterious minor alleles divided by the individual number of polymorphic sites. **H**: Count of homozygous highly deleterious minor alleles divided by the individual number of polymorphic sites.

This plot shows the individual count of minor alleles divided by the individual number of polymorphic sites per variants category (A: Neutral; B: Lowly deleterious, C: Moderately deleterious; D: Highly deleterious; E: homozygous neutral; F: Homozygous lowly deleterious, G: Homozygous moderately deleterious; H: Homozygous highly deleterious) in refugium populations versus recolonized populations. There is no significant difference between both groups of populations for all types of minor alleles both for single copy and homozygous states (Neutral single copy: $W = 8125$, $p\text{-value} = 0.9382$; Lowly deleterious single copy: $W = 8161$, $p\text{-value} = 0.9762$; Mildly deleterious single copy: $W = 8383$, $p\text{-value} = 0.7924$; Highly deleterious single copy: $W = 6857$, $p\text{-value} = 0.0786$; Neutral homozygous: $W = 8518$, $p\text{-value} = 0.6583$; Lowly deleterious single copy: $W = 8334$, $p\text{-value} = 0.8429$; Mildly deleterious single copy: $W = 8795$, $p\text{-value} = 0.4182$; Highly deleterious single copy: $W = 7331$, $p\text{-value} = 0.2583$).

Comparing populations in details

When we zoom into the population-specific F_{HBD} (Figure S2), mean sum of HBD lengths per HBD class (figure S6), and N_e estimation (table 1), the pattern differs among populations. Concerning islands populations, both the EC and WC populations displayed high F_{HBD} (mean $F_{HBD} = 0.132$ and 0.112 for EC and WC respectively, with large variance among individuals for the EC though), were enriched in HBD classes from 32g to 128g ago. Additionally, their N_e estimation was higher than the other islands. We found that the AE population had low HBD coefficients (mean $F_{HBD} = 0.061$), very similar to those from GR, was slightly enriched in HBD segments coalescing 32g and 64g ago, and had very high N_e estimation, which were very similar to GR. The distribution of HBD segments in the IO population was very similar to that of the AE populations, but they had slightly longer sums of lengths for segments coalescing 256g ago. Additionally, individuals from the IO population displayed slightly higher F_{HBD} values (mean $F_{HBD} = 0.081$) and lower N_e estimates. In the CO population, F_{HBD} coefficients were high (mean $F_{HBD} = 0.098$) and the HBD segments distribution was highly enriched in segments coalescing 256 generations ago. In addition, the N_e estimation for the CO population was among the lowest. From all the populations studied, CT appeared to be one of the most inbred with

low N_e , high F_{HBD} values (mean $F_{HBD} = 0.134$) and a substantial percentage of HBD segments from 128 generations ago. Compared to other islands, the CY population had lowest F_{HBD} coefficients (mean $F_{HBD} = 0.073$) and was enriched in HBD segments from 64g and 128g ago. In addition, the CY N_e estimate was relatively high compared to those of the other islands. Finally, the IR and GB populations had high F_{HBD} (mean $F_{HBD} = 0.104$ and 0.102 for GB and IR, respectively), the lowest N_e estimation, and were enriched for HBD segments dating back to 256g and 512g.

Concerning continental populations, they displayed lower F_{HBD} and higher N_e estimation. CH and DK individuals displayed low inbreeding coefficient. Concerning continental populations, they displayed lower F_{HBD} and higher N_e estimation. CH and DK individuals displayed low inbreeding coefficients (mean $F_{HBD} = 0.040$ and 0.041 respectively), high N_e estimations, and were highly enriched in HBD segments that coalesced over 512 million years ago. In contrast to the other continental populations, the FR population had high values of F_{HBD} (mean $F_{HBD} = 0.092$) and was enriched in HBD segments that coalesced 16g, 32g, 256g and 512g ago. It showed, however, high N_e estimates comparable to those of the rest of the continental populations. In the IS, PT, and MA populations, there were low F_{HBD} values (mean F_{HBD} values are 0.042, 0.024, and 0.011 for IS, PT, and MA, respectively), no enrichment in any HBD segment class, and a high estimated N_e value. The mean F_{HBD} of the GE population was the highest among continental populations (mean $F_{HBD} = 0.069$). Furthermore, compared with other continental populations, their sum of HBD segment lengths was also enriched for segments that coalesced between 64g and 128g ago (although they were still smaller than most island populations). In addition, they displayed an extremely low estimated N_e . The IT and SB populations had a mean F_{HBD} of 0.046 and 0.054, respectively, and did not show any enrichment in any HBD segment classes. Similarly, their N_e estimation was very similar to the other recolonized continental populations. Finally, the GR population showed an average inbreeding coefficient value and a mean sum of HBD segment lengths (mean $F_{HBD} = 0.060$). In addition, the GR N_e estimate was relatively large compared with other recolonized continental populations.

The AE displayed the highest N_e estimation, very close to the one from GR. This is concordant with previous studies which showed that there is low F_{ST} between these two populations [1, 3]. Concerning the IO population, the N_e estimation was slightly lower compared to the AE population (but higher than GB, IR, CO and CT) and HBD segments

numbers and lengths were very similar to other island populations. Historically this population has been shown to be genetically close to GR as well [3]. The CY population also showed high N_e estimation and low F_{HBD} distribution (compared to other islands). In addition, the number of HBD segments were the lowest among the island populations. The high diversity we found is consistent with previous studies and might be explained because this population is genetically close to the IS population [3, 1]. The populations from CT and CO displayed even lower N_e and higher numbers of HBD segments. CT especially showed inflated F_{HBD} distribution. This is consistent with its small size and stronger isolation from the rest of the populations [3]. On the contrary, CO showed especially high number of small HBD segments suggesting no recent mating between closely related individuals but an history of small N_e and long-term isolation. The Canary islands populations (EC and WC) showed N_e estimation close other small islands namely CT and CO and relatively high F_{HBD} estimation, especially with one inbred individual in the EC population. Similarly, their sum of HBD segments lengths were clustering with the other island populations. All these are consistent with their history of old colonisation and long-term isolation [2]. Populations from DK, IT and SB displayed N_e , F_{HBD} estimation and total numbers of HBD segments close to CH. This is not surprising as we know that there is shallow population differentiation between these populations (especially between CH and DK) [1]. The population from GR displayed similar statistics compared to the rest of central Europe (namely CH, FR, DK, IT and SB) and had very similar estimates to the AE population. Concerning the MA population, we estimated a high N_e and very low inbreeding coefficients. This is concordant with previous studies which found that this population is close to the PT population [2] known as the biggest population and a refugium during the last glacial maxima [1]. This is consistent with our results as it harbored the highest N_e and low F_{HBD} in our analyses. In the MA population, we observed three individuals below the line in the F_{AS} vs F_{HBD} plot. These individuals are from the same family. Similar to what has been observed for the CH individuals, these individuals are shifted towards the line when we trim the data set for relatedness. Concerning the population from IS, we estimated a high N_e and a F_{HBD} distribution as well as total number of HBD segments similar to what was observed for the CH, FR and DK populations. The high N_e is concordant with previous studies who found that the IS population is a reservoir of diversity [1, 3].

References

- [1] T. Cumer, A.P. Machado, G. Dumont, V. Bontzorlos, R. Ceccherelli, M. Charter, K. Dichmann, N. Kassinis, R. Lourenço, F. Manzia, H-D. Martens, L. Prévost, M. Rakovic, I. Roque, F. Siverio, A. Roulin, and J. Goudet. Landscape and Climatic Variations Shaped Secondary Contacts amid Barn Owls of the Western Palearctic. *Molecular Biology and Evolution*, 39(1):msab343, 2022.
- [2] T. Cumer, A.P. Machado, F. Siverio, S.I. Cherkaoui, I. Roque, R. Lourenço, M. Charter, A. Roulin, and J. Goudet. Genomic basis of insularity and ecological divergence in barn owls (*Tyto alba*) of the Canary Islands. *Heredity*, 129(5):281–294, 2022.
- [3] A.P. Machado, A. Topaloudis, T. Cumer, E. Lavanchy, V. Bontzorlos, R. Ceccherelli, M. Charter, N. Kassinis, P. Lymberakis, F. Manzia, A-L. Ducrest, M. Dupasquier, N. Guex, A. Roulin, and J. Goudet. Genomic consequences of colonisation, migration and genetic drift in barn owl insular populations of the eastern Mediterranean. *Molecular Ecology*, 31(5):1375–1388, 2022.

Supplementary Material Chapter IV: Inbreeding depression in the Swiss barn owl (*Tyto alba*) population

Authors: Eléonore Lavanchy, Anna Hewett, Tristan Cumer, Alexandros Topaloudis, Anne-Lyse Ducrest, Céline Simon, Alexandre Roulin and Jérôme Goudet

Supplementary Material & Methods

As mentioned in the material and methods section, we did not describe and present all the models in the main text. The following equations describe all the animal models used in this paper. Age (in days) for juveniles was modeled as a Gompertz function with parameters estimated from the full dataset.

$$\begin{aligned} \text{Bill Length juveniles} &= \beta_0 + \beta_1 F_{uni}^W + \beta_2 \text{Sex} \\ &+ \beta_3 \text{rank} + \beta_4 (179.8 * e^{-1.127 * 0.921^{age}}) + \alpha^{\text{pedigree}} \\ &+ \alpha^{\text{individual}} + \alpha^{\text{observer}} + \alpha^{\text{clutch}} + \alpha^{\text{year}} \end{aligned} \quad (\text{S1})$$

$$\begin{aligned} \text{Bill Length adults} &= \beta_0 + \beta_1 \text{Sex} \\ &+ \alpha^{\text{pedigree}} + \alpha^{\text{individual}} + \alpha^{\text{observer}} + \alpha^{\text{clutch}} + \alpha^{\text{year}} \end{aligned} \quad (\text{S2})$$

$$\begin{aligned} \text{Mass juveniles} &= \beta_0 + \beta_1 F_{uni}^W + \beta_2 \text{Sex} \\ &+ \beta_3 \text{rank} + \beta_4 (369.5 * e^{-4.425 * 0.891^{age}}) + \alpha^{\text{pedigree}} \\ &+ \alpha^{\text{individual}} + \alpha^{\text{observer}} + \alpha^{\text{clutch}} + \alpha^{\text{year}} \end{aligned} \quad (\text{S3})$$

$$\begin{aligned} \text{Mass adults} &= \beta_0 + \beta_1 \text{Sex} * \beta_2 \text{lastChickHatched} \\ &+ \alpha^{\text{pedigree}} + \alpha^{\text{individual}} + \alpha^{\text{observer}} + \alpha^{\text{clutch}} + \alpha^{\text{year}} \end{aligned} \quad (\text{S4})$$

$$\begin{aligned} \text{Tarsus Length juveniles} &= \beta_0 + \beta_1 F_{uni}^W \\ &+ \beta_2 \text{Sex} + \beta_3 \text{rank} + \beta_4 (719.5 * e^{-2.274 * 0.890^{age}}) \\ &+ \alpha^{\text{pedigree}} + \alpha^{\text{individual}} + \alpha^{\text{observer}} + \alpha^{\text{clutch}} + \alpha^{\text{year}} \end{aligned} \quad (\text{S5})$$

$$\begin{aligned} \text{Tarsus Length adults} &= \beta_0 + \beta_1 \text{Sex} + \beta_2 \text{rank} \\ &+ \alpha^{\text{pedigree}} + \alpha^{\text{individual}} + \alpha^{\text{observer}} + \alpha^{\text{clutch}} + \alpha^{\text{year}} \end{aligned} \quad (\text{S6})$$

$$\begin{aligned}
& \text{Bill Length juveniles} = \beta_0 + \beta_1 F_{HBD} \\
& + \beta_2 \text{Sex} + \beta_3 \text{rank} + \beta_4 (179.8 * e^{-1.127 * 0.921^{age}}) \\
& + \alpha^{pedigree} + \alpha^{individual} + \alpha^{observer} + \alpha^{clutch} + \alpha^{year}
\end{aligned} \tag{S7}$$

$$\begin{aligned}
& \text{Bill Length adults} = \beta_0 + \beta_1 F_{uni}^W + \beta_2 \text{Sex} \\
& + \alpha^{pedigree} + \alpha^{individual} + \alpha^{observer} + \alpha^{clutch} + \alpha^{year}
\end{aligned} \tag{S8}$$

$$\begin{aligned}
& \text{Bill Length adults} = \beta_0 + \beta_1 F_{HBD} + \beta_2 \text{Sex} \\
& + \alpha^{pedigree} + \alpha^{individual} + \alpha^{observer} + \alpha^{clutch} + \alpha^{year}
\end{aligned} \tag{S9}$$

$$\begin{aligned}
& \text{Mass juveniles} = \beta_0 + \beta_1 F_{HBD} + \beta_2 \text{Sex} \\
& + \beta_3 \text{rank} + \beta_4 (369.5 * e^{-4.425 * 0.891^{age}}) + \alpha^{pedigree} \\
& + \alpha^{individual} + \alpha^{observer} + \alpha^{clutch} + \alpha^{year}
\end{aligned} \tag{S10}$$

$$\begin{aligned}
& \text{Mass adults} = \beta_0 + \beta_1 F_{uni}^W + \beta_2 \text{Sex} * \beta_3 \text{lastChickHatched} \\
& + \alpha^{pedigree} + \alpha^{individual} + \alpha^{observer} + \alpha^{clutch} + \alpha^{year}
\end{aligned} \tag{S11}$$

$$\begin{aligned}
& \text{Mass adults} = \beta_0 + \beta_1 F_{HBD} + \beta_2 \text{Sex} * \beta_3 \text{lastChickHatched} \\
& + \alpha^{pedigree} + \alpha^{individual} + \alpha^{observer} + \alpha^{clutch} + \alpha^{year}
\end{aligned} \tag{S12}$$

$$\begin{aligned}
& \text{Tarsus Length juveniles} = \beta_0 + \beta_1 F_{HBD} \\
& + \beta_2 \text{Sex} + \beta_3 \text{rank} + \beta_4 (719.5 * e^{-2.274 * 0.890^{age}}) \\
& + \alpha^{pedigree} + \alpha^{individual} + \alpha^{observer} + \alpha^{clutch} + \alpha^{year}
\end{aligned} \tag{S13}$$

$$\begin{aligned}
& \text{Tarsus Length adults} = \beta_0 + \beta_1 \text{Sex} + \beta_2 F_{UNI}^W + \beta_3 \text{rank} \\
& + \alpha^{pedigree} + \alpha^{individual} + \alpha^{observer} + \alpha^{clutch} + \alpha^{year}
\end{aligned} \tag{S14}$$

$$\begin{aligned}
& \text{Tarsus Length adults} = \beta_0 + \beta_1 \text{Sex} + \beta_2 F_{HBD} + \beta_3 \text{rank} \\
& + \alpha^{pedigree} + \alpha^{individual} + \alpha^{observer} + \alpha^{clutch} + \alpha^{year}
\end{aligned} \tag{S15}$$

We also ran inbreeding depression models without the pedigree and with both inbreeding coefficients:

$$\begin{aligned}
\text{Bill Length juveniles} &= \beta_0 + \beta_1 F_{uni}^W + \beta_2 \text{Sex} \\
&+ \beta_3 \text{rank} + \beta_4 (179.8 * e^{-1.127 * 0.921^{age}}) \\
&+ \alpha^{individual} + \alpha^{observer} + \alpha^{clutch} + \alpha^{year}
\end{aligned} \tag{S16}$$

$$\begin{aligned}
\text{Bill Length juveniles} &= \beta_0 + \beta_1 F_{HBD} + \beta_2 \text{Sex} \\
&+ \beta_3 \text{rank} + \beta_4 (179.8 * e^{-1.127 * 0.921^{age}}) \\
&+ \alpha^{individual} + \alpha^{observer} + \alpha^{clutch} + \alpha^{year}
\end{aligned} \tag{S17}$$

$$\begin{aligned}
\text{Bill Length adults} &= \beta_0 + \beta_1 F_{UNI}^W + \beta_2 \text{Sex} + \beta_3 \text{rank} \\
&+ \alpha^{individual} + \alpha^{observer} + \alpha^{clutch} + \alpha^{year}
\end{aligned} \tag{S18}$$

$$\begin{aligned}
\text{Bill Length adults} &= \beta_0 + \beta_1 F_{HBD} + \beta_2 \text{Sex} + \beta_3 \text{rank} \\
&+ \alpha^{individual} + \alpha^{observer} + \alpha^{clutch} + \alpha^{year}
\end{aligned} \tag{S19}$$

$$\begin{aligned}
\text{Mass juveniles} &= \beta_0 + \beta_1 F_{UNI}^W + \beta_2 \text{Sex} \\
&+ \beta_3 \text{rank} + \beta_4 (369.5 * e^{-4.425 * 0.891^{age}}) \\
&+ \alpha^{individual} + \alpha^{observer} + \alpha^{clutch} + \alpha^{year}
\end{aligned} \tag{S20}$$

$$\begin{aligned}
\text{Mass juveniles} &= \beta_0 + \beta_1 F_{HBD} + \beta_2 \text{Sex} \\
&+ \beta_3 \text{rank} + \beta_4 (369.5 * e^{-4.425 * 0.891^{age}}) \\
&+ \alpha^{individual} + \alpha^{observer} + \alpha^{clutch} + \alpha^{year}
\end{aligned} \tag{S21}$$

$$\begin{aligned}
\text{Mass adults} &= \beta_0 + \beta_1 F_{uni}^W + \beta_2 \text{Sex} * \beta_3 \text{lastChickHatched} \\
&+ \alpha^{individual} + \alpha^{observer} + \alpha^{clutch} + \alpha^{year}
\end{aligned} \tag{S22}$$

$$\begin{aligned}
\text{Mass adults} &= \beta_0 + \beta_1 F_{UNI}^W + \beta_2 \text{Sex} * \beta_3 \text{lastChickHatched} \\
&+ \alpha^{individual} + \alpha^{observer} + \alpha^{clutch} + \alpha^{year}
\end{aligned} \tag{S23}$$

$$\begin{aligned}
\text{Tarsus Length juveniles} &= \beta_0 + \beta_1 F_{UNI}^W \\
&+ \beta_2 \text{Sex} + \beta_3 \text{rank} + \beta_4 (719.5 * e^{-2.274 * 0.890^{age}}) \\
&+ \alpha^{individual} + \alpha^{observer} + \alpha^{clutch} + \alpha^{year}
\end{aligned} \tag{S24}$$

$$\begin{aligned}
& \textit{Tarsus Length juveniles} = \beta_0 + \beta_1 F_{HBD} \\
& + \beta_2 \textit{Sex} + \beta_3 \textit{rank} + \beta_4 (719.5 * e^{-2.274 * 0.890^{age}}) \\
& + \alpha^{individual} + \alpha^{observer} + \alpha^{clutch} + \alpha^{year}
\end{aligned} \tag{S25}$$

$$\begin{aligned}
& \textit{Tarsus Length adults} = \beta_0 + \beta_1 \textit{Sex} + \beta_2 F_{UNI}^W \\
& + \beta_3 \textit{rank} + \alpha^{individual} + \alpha^{observer} + \alpha^{clutch} + \alpha^{year}
\end{aligned} \tag{S26}$$

$$\begin{aligned}
& \textit{Tarsus Length adults} = \beta_0 + \beta_1 \textit{Sex} + \beta_2 F_{HBD} \\
& + \beta_3 \textit{rank} + \alpha^{individual} + \alpha^{observer} + \alpha^{clutch} + \alpha^{year}
\end{aligned} \tag{S27}$$

For the yearly fitness-related traits we only built models for adults and only for females for the number of eggs laid:

$$\begin{aligned}
& \# \textit{eggs laid adults} = \beta_0 + \beta_1 \textit{Female} + \beta_2 F_{UNI}^W \\
& + \beta_3 \textit{Tarsus length} + \beta_4 \textit{Julian day} + \beta_5 \textit{Female age} \\
& + \alpha^{individual} + \alpha^{Site ID} + \alpha^{Year}
\end{aligned} \tag{S28}$$

$$\begin{aligned}
& \# \textit{eggs laid adults} = \beta_0 + \beta_1 \textit{Female} + \beta_2 F_{HBD} \\
& + \beta_3 \textit{Tarsus length} + \beta_4 \textit{Julian day} + \beta_5 \textit{Female age} \\
& + \alpha^{individual} + \alpha^{Site ID} + \alpha^{Year}
\end{aligned} \tag{S29}$$

$$\begin{aligned}
& \# \textit{eggs laid adults} = \beta_0 + \beta_1 \textit{Female} + \beta_2 F_{UNI}^W \\
& + \beta_3 \textit{Tarsus length} + \beta_4 \textit{Julian day} + \beta_5 \textit{Female age} \\
& + \alpha^{pedigree} + \alpha^{individual} + \alpha^{Site ID} + \alpha^{Year}
\end{aligned} \tag{S30}$$

$$\begin{aligned}
& \# \textit{eggs laid adults} = \beta_0 + \beta_1 \textit{Female} + \beta_2 F_{HBD} \\
& + \beta_3 \textit{Tarsus length} + \beta_4 \textit{Julian day} + \beta_5 \textit{Female age} \\
& + \alpha^{pedigree} + \alpha^{individual} + \alpha^{Site ID} + \alpha^{Year}
\end{aligned} \tag{S31}$$

$$\begin{aligned}
& \textit{Pr}(\textit{hatching} = 1) = \textit{probit}^{-1}(\beta_0 + \beta_1 \textit{Julian day} \\
& + \beta_3 \textit{Female} F_{uni}^W + \beta_4 \textit{Male} F_{uni}^W + \alpha^{clutch})
\end{aligned} \tag{S32}$$

$$\begin{aligned}
& \textit{Pr}(\textit{hatching} = 1) = \textit{probit}^{-1}(\beta_0 + \beta_1 \textit{Julian day} \\
& + \beta_3 \textit{Female} F_{HBD} + \beta_4 \textit{Male} F_{HBD} + \alpha^{clutch})
\end{aligned} \tag{S33}$$

Supplementary Results

Figure S1 shows the different heritability (h^2) estimates distribution (among the MCMC chain iterations) for the three morphological traits for both adults and juveniles. Heritability was always much higher for adults compared to juveniles and was the lowest for mass.

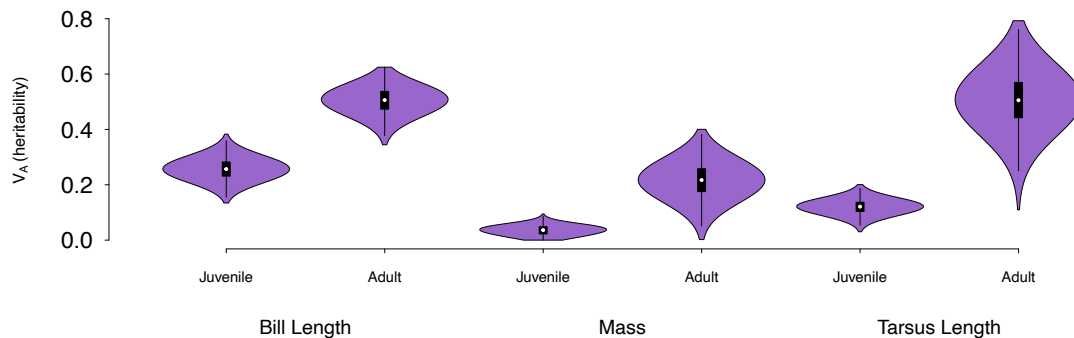


Figure S1 : h^2 estimates distribution (among the MCMC chain iterations) for the three morphological traits (bill length, mass and tarsus length) for adults and juveniles. Juveniles' models included the inbreeding coefficient F_{UNI} as a fixed effect but this was not the case for adults.

Figure S2 shows the different β estimates for all the models we ran to quantify inbreeding depression for the five phenotypes: morphological traits bill length, mass and tarsus length and yearly fitness-related traits number of eggs laid (for females only) as well as probability that an egg hatches according to both parents inbreeding coefficient. For morphological traits, we compare juveniles and adults, F_{UNI} and F_{HBD} as well as including the pedigree in the model or not. For yearly fitness-related traits, we compare F_{HBD} and F_{UNI} as well as including the pedigree in the model versus the simple model for number of eggs laid in adults only. It is striking to note that there is little difference between the models including or excluding the pedigree. This is not surprising since our data does not contain a strong structure (although we do have family structure). The second striking result is that the CI for β estimated with F_{HBD} are always larger than the estimates from F_{UNI} and this is especially true for the binomial models (modeling the probability that an egg hatches according to the parents inbreeding coefficients, panel C) and to a lesser extent for adult morphological traits (panel B). Despite the fact that we cannot know the true value of β here, the smaller CI around β , coupled with the fact that the owls we are studying come from a population with large effect sizes and numbers, suggests that F_{UNI} is likely more accurate than F_{HBD} [1, 2].

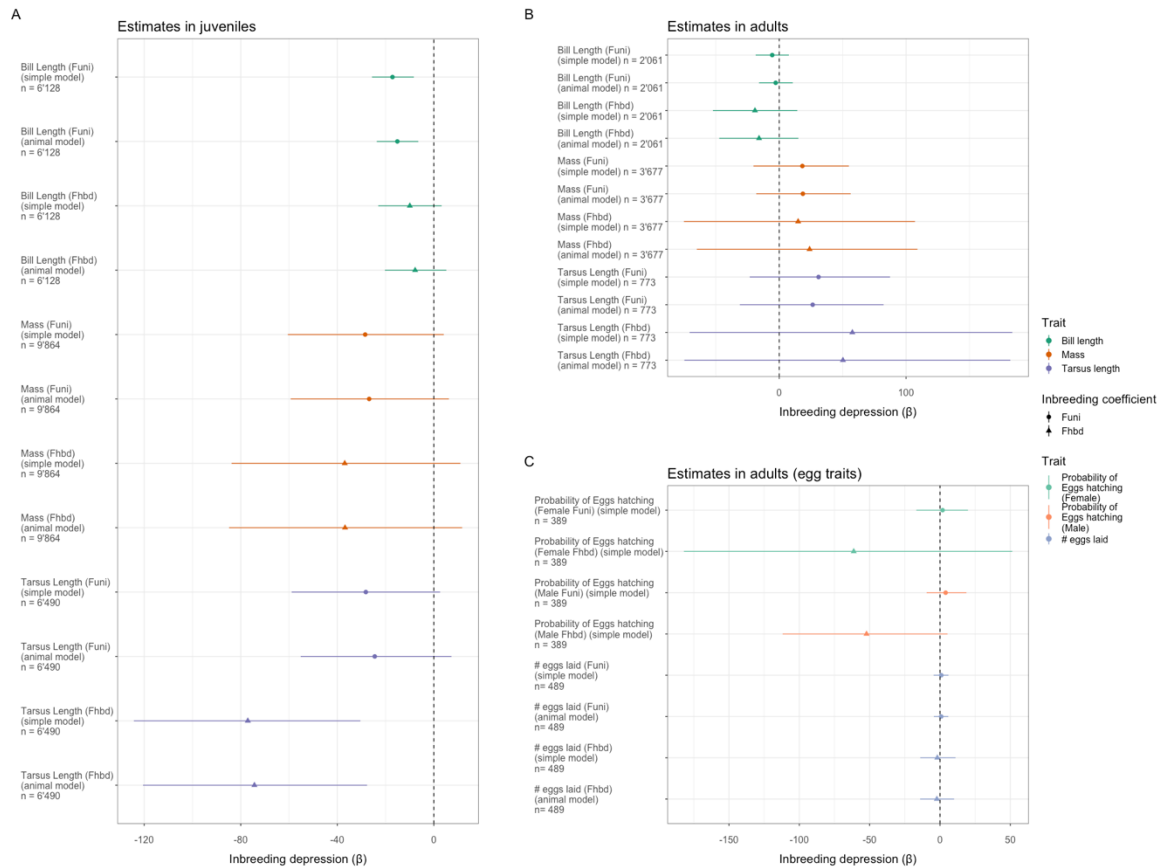


Figure S2: β estimates for all the inbreeding depression models **A:** β estimates in juveniles for the three morphological traits (bill length, mass and tarsus length) for both inbreeding coefficients: F_{UNI} and F_{HBD} and with and without including the pedigree as a random factor. Models correspond to equations S16, S1, S17, S7, S20, S3, S21, S10, S24, S5, S25 and S13. **B:** β estimates in adults for the three morphological traits (bill length, mass and tarsus length) for both inbreeding coefficients: F_{UNI} and F_{HBD} and with inbreeding depression model (with and without the pedigree). Models correspond to equations S18, S8, S19, S9, S22, S11, S23, S12, S26, S14, S27 and S15. **C:** β estimates in adults for the two fitness-related traits (number of eggs laid (for females only and with both inbreeding depression models) and probability that an eggs hatches (for both sexes, both inbreeding coefficients and with the simple inbreeding depression model only). These models correspond to equations S32, S33, S28, S30, S29 and S31.

Figure S3 shows the difference between inbreeding depression models when the rank of the individual (in birth) is included or not for tarsus length in adults. Rank had a significant effect on tarsus length so we decided to include it in the model in the main text. However, this greatly reduced the sample size as well as the range of F_{UNI} covered and changed the sign of the β estimate. Ideally, we would like to be able to include rank and increase our range of F_{UNI} , however this is not possible with our data. The model including rank might be biased because the range of F_{UNI} we cover is very small but the model without including rank might be biased because we lack an important covariate. Overall, it is not of striking importance since both models yield non-significant β but I would say I do not trust any of the model for reliable conclusion on the effect of inbreeding on tarsus length in adults.

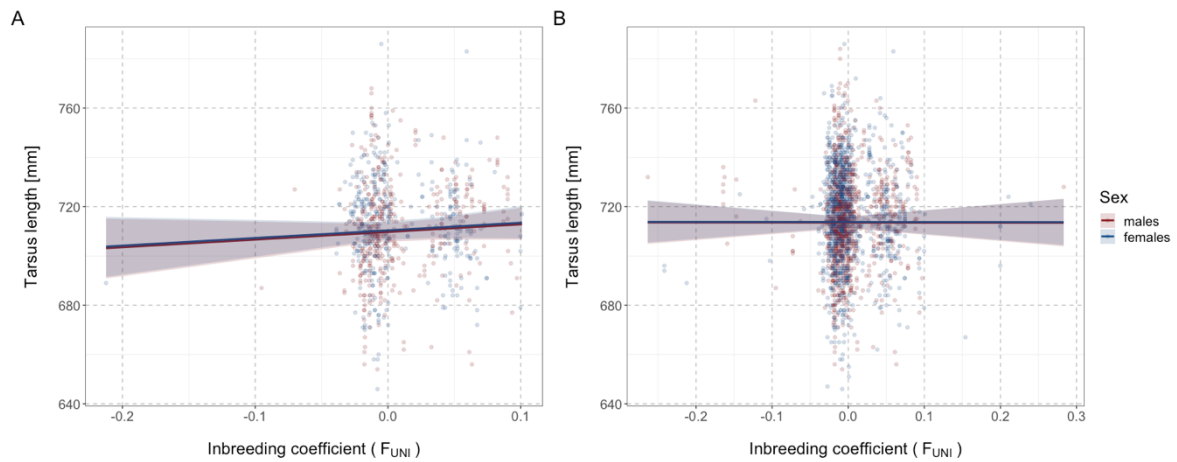


Figure S3: Regression plot of tarsus length according to F_{UNI} in adults. **A:** The model does include the rank of the individual at birth as a covariate. **B:** the model does not include rank of the individual at birth as a covariate.

Figure S4 shows the distributions of F_{UNI} in juveniles and adults. Both distributions are significantly different (Wilcoxon rank test; $W = 1421740$, p -value = 0.001755) but what is really interesting is that we have more individuals with $F_{UNI} > 0.1$ in juveniles compared to adults.

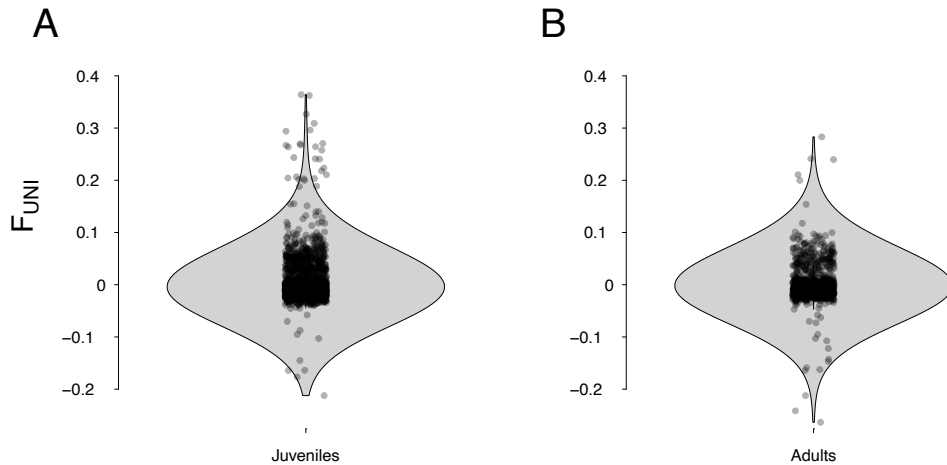


Figure S4: F_{UNI} distributions in juveniles (**A**) and adults (**B**).

References

- [1] S.W. Alemu, N.K. Kadri, C. Harland, P. Faux, C. Charlier, A. Caballero, and T. Druet. An evaluation of inbreeding measures using a whole-genome sequenced cattle pedigree. *Heredity*, 126(3):410–423, 2021.
- [2] A. Caballero, B. Villanueva, and T. Druet. On the estimation of inbreeding depression using different measures of inbreeding from molecular markers. *Evolutionary Applications*, 14(2):416–428, 2021.

ANNEX I

Authors : Alexandros Topaloudis^{1,2}, Eléonore Lavanchy^{1,2}, Tristan Cumer^{1,2}, Anne-Lyse Ducrest¹, Céline Simon¹, Alexandre Roulin^{1,2}, Jérôme Goudet^{1,2}

¹ Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland

² Swiss Institute of Bioinformatics, University of Lausanne, Lausanne, Switzerland

Presented here is a paper in which I participated, but did not write. This document is included here as an annex since it describes the SNP calling of the 502 owls referred to as the “Reference panel” as well as the recombination maps used in chapters III and IV. We could not cite this paper because it is not published (or submitted as a preprint) yet.

My contribution to this paper was the SNP calling.

Supplementary material PDF for this paper can be found on [GitHub](#).

The recombination landscape of the barn owl, from families to populations

Alexandros Topaloudis^{1,2} - 0000-0002-0909-8695

Eleonore Lavanchy^{1,2} - 0000-0003-4951-9332

Tristan Cumer^{1,2} - 0000-0002-0276-7462

Anne-Lyse Ducrest¹ - 0000-0001-6412-2769

Celine Simon ¹

Nika Paposhvili³ - 0000-0002-3785-6174

Alexandre Roulin¹ - 0000-0003-1940-6927

Jerome Goudet^{1,2} - 0000-0002-5318-7601

Affiliations:

1 : Department of Ecology and Evolution, University of Lausanne, Lausanne 1011, Switzerland

2 : Swiss Institute of Bioinformatics, Lausanne 1011, Switzerland

3 : Ilia State University, Tbilisi 0162, Georgia

Correspondence:

Alexandros Topaloudis, Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland. Email: alexandros.topaloudis@unil.ch

Jerome Goudet, Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland. Email: jerome.goudet@unil.ch

Keywords:

Linkage mapping, LD, meiotic recombination, heterochiasmy, hotspots

Abstract

Homologous recombination is a meiotic process that generates diversity along the genome and interacts with all evolutionary forces. Despite its importance, studies of recombination landscapes are lacking due to methodological limitations and a dearth of appropriate data. In fact, different methods of inference overcome different problems of inference. Linkage mapping of familial data gives unbiased broad-scale estimates of recombination and LD-based inference provides fine-scale data albeit depended on the effective population size and acting selective forces. In this study we use an intersection of methods, on a family dataset of whole genome sequences and elucidate the first owl recombination landscape for the barn owl (*Tyto alba*). Using linkage mapping we refine the genome assembly to a chromosome-level quality, and identify subtle heterochiasmy through fine-scale differences in crossover placement and shuffling proportions between male and female barn owls. Additionally, through fine-scale LD-based inference we show that the abundant variation in recombination landscapes of different linkage groups is guided by their length and strongly shapes the genetic diversity of the genome. Furthermore, we identify recombination hotspots showing little evolutionary stability of the fine-scale recombination landscape through population-level comparisons. Overall, this comprehensive analysis enhances our understanding of recombination dynamics, genomic architecture, and sex-specific variation in the barn owl, contributing valuable insights to the broader field of avian genomics.

Introduction

Recombination is a key feature of sexual reproduction that has multiple evolutionary implications but its inference is often overlooked in non-model species. Homologous meiotic recombination, hereafter recombination, is the reciprocal exchange of genetic material between homologous chromosomes during the first meiotic division. The physical exchange, called a crossover, generates the necessary tension between chromosomes to ensure their

proper segregation in the daughter cells and its absence has been associated with aneuploidy (Hassold et al., 2007; Koehler et al., 1996; Zelkowski et al., 2019; Zickler & Kleckner, 2015).

Beyond contributing to the integrity of proper meiotic division, recombination can also have evolutionary consequences by shuffling alleles between haplotypes thus affecting the genomic composition of a population. This shuffling brings about evolutionary benefits including faster adaptation to a changing environment and more efficient selection (Hill & Robertson, 1966; Otto & Lenormand, 2002) . Despite its benefits, recombination can also impede adaptation by breaking up beneficial combinations of alleles or increasing the rate of mutations and chromosomal rearrangements (Arbeithuber et al., 2015; Barton & Charlesworth, 1998; Halldorsson et al., 2019). Finally, because recombination rates vary along the genomic sequence, they affect almost all genome-wide analyses. For instance, genetic diversity along the genome will correlate with recombination rates due to the effect of linked selection (Begun & Aquadro, 1992) and regions of low recombination can appear as false positives in scans for selection based on differentiation (Booker et al., 2020). It is thus important to know the position of crossovers in the genome and the frequency with which they occur to be able to account for recombination variation and eliminate confounding with other evolutionary forces.

However, quantifying recombination is a laborious task. One approach is linkage mapping, the positioning of markers along the sequence with a distance proportional to the recombination rate between them. This approach requires family data (or when available, controlled crosses) and has been applied to several species so far providing both a reliable measure of crossing over frequency (Brazier & Glémin, 2022; Kong et al., 2002; Stapley et al., 2017). Linkage mapping can also quantify the differences in recombination rates between sexes (i.e. heterochiasmy, Brekke et al., 2022; Johnston et al., 2017; Kong et al., 2010). Unfortunately, linkage mapping requires family data which are only available in a few study species (Peñalba & Wolf, 2020). Further, recombination rates estimated with this

method are limited by the number of meioses observed, making it impossible to quantify it accurately in small genomic windows with realistic sample sizes (Halldorsson et al., 2019).

To address this problem, additional approaches have been developed to estimate recombination using whole genome sequences of tens of unrelated individuals (Auton & McVean, 2007; Chan et al., 2012; Spence & Song, 2019). These methods model the observed linkage disequilibrium (LD) between markers, assessing ancestral recombination events that occurred in the coalescent history of the sample (N. Li & Stephens, 2003). This approach, hereafter LD-based inference, has enabled the quantification of fine scale recombination variation initially in humans (McVean et al., 2004; Myers et al., 2005). However, because of the limited genomic resources required, LD-based inference of recombination has also been applied to non-model species like birds, reptiles and fish among others (ex. Kawakami et al., 2017; Schield et al., 2020; Shanfelter et al., 2019; Singhal et al., 2015). Such fine-scale inferences have implicated the PRDM9 gene for the location and evolutionary turnover of recombination hotspots (narrow regions of increased recombination) of some species (Booker et al., 2017; Myers et al., 2010). Similarly LD-based inference has shown that species that lack the PRDM9 gene show evolutionary conserved hotspots in regions of accessible chromatin (Auton et al., 2013; Baker et al., 2017; Singhal et al., 2015), or no hotspots at all (Kaur & Rockman, 2014; Smukowski Heil et al., 2015).

Despite the discoveries and benefits of quantifying fine-scale variation with few genomes, LD-based inference has certain limitations. The method infers the population recombination rate (ρ), the product of the effective population size (N_e) and the recombination rate, a caveat that has two major implications. It does not distinguish between crossing over in male and female meioses and is affected by forces that change N_e and not the recombination rate itself. This essentially means that selection can confound estimates of recombination (O'Reilly et al., 2008). To remedy this and while selection remains a confounding factor, accounting for demography has been implemented in recent applications (Spence & Song, 2019). Even then, estimates of LD-based methods are often validated with a different

estimate of recombination, for example linkage mapping (Axelsson et al., 2012; McVean et al., 2004; Shanfelter et al., 2019; Wall et al., 2022).

Therefore, a combination of approaches is the preferred route to accurately infer the recombination rates in the fine scale. However, such combinations will require both whole genome sequencing data of unrelated individuals and family data, a doubly costly resource. Especially since family data, beyond controlled crosses, are usually only available for specific species. Family data may be disproportionately available for wild populations of birds because their nesting behaviour can facilitate population monitoring and can be exploited for the construction of long term pedigrees (Grant & Grant, 2002; Lack & Lack, 1958; Pemberton, 2008). Despite this opportunity for the generation of family data in birds, we only have information about the recombination landscapes of very few genera. Most information for the order comes from studies of pedigreed populations using linkage mapping (Backström et al., 2010; Groenen et al., 2009; Hagen et al., 2020; Kawakami et al., 2014; Peñalba et al., 2020; van Oers et al., 2014) and only two species benefit from both an LD-based inference and a linkage-mapping approach: the zebra finch (*Taeniopygia guttata*) and the collared flycatcher (*Ficedula albicollis*) (Kawakami et al., 2017; Singhal et al., 2015). All these previous studies have shown that recombination in birds exhibits broad-scale among-species variation, inconclusive patterns of sex-differences and the absence of the PRDM9 gene. Firstly, rates of recombination inferred from linkage mapping tend to differ between species despite a rather conserved avian karyotype. For example two members of the passerine clade (collared flycatcher and the superb fairy-wren *Malurus cyaneus*) show an unexplained twofold difference in genetic length among their largest syntenic chromosomes (Kawakami et al., 2014; Peñalba et al., 2020). Further, birds show evidence for lack of consistent patterns of heterochiasmy (Sardell & Kirkpatrick, 2019). Recent studies (McAuley et al., 2023; Zhang et al., 2023) point at evidence for fine-scale differences in the placement of crossovers between sexes, an observation overlooked when focusing on total genetic lengths. Finally, birds lack the PRDM9 hotspot gene (Baker et al., 2017). While the evolutionary stability of hotspots has been demonstrated in finches and flycatchers

(Kawakami et al., 2017; Singhal et al., 2015), it is still unknown how to reconcile such findings with evidence for broad scale variation. Importantly, the bird species for which recombination rate variation has been studied so far all belong to the passerine order apart from the economically important chicken (*Gallus gallus*). Therefore information from other avian orders is lacking and the generality of such findings can be questioned.

Here, using both linkage mapping and LD-information we present the first recombination landscape for a species of the owl order, the barn owl (*Tyto alba*). We use this species because it has the highest quality genome assembly of an owl species (Ducrest et al., 2020; Machado, Cumer, et al., 2022), a set of whole-genome sequences available from past studies (Cumer, Machado, Dumont, et al., 2022a; Cumer, Machado, Siverio, et al., 2022; Cumer et al., 2024; Machado, Cumer, et al., 2022; Machado, Topaloudis, et al., 2022) and a long-term pedigreed population with an untapped genomic potential (Charmantier et al., 2014; Sheldon et al., 2022). We capitalise on 176 genomes previously published along with 326 newly sequenced to build a high confidence variant set that spans the diversity of the species across the Western palearctic. Specifically, we used linkage mapping on a dataset of 250 owls belonging to 28 families to identify linkage groups in the barn owl sequence assembly, estimate the sex-averaged linkage map length and quantify sex-differences in recombination. Additionally, we used an LD-based approach on 102 unrelated individuals from three populations to infer fine-scale recombination rate variation and scale our results using the estimates from the linkage map. Using these complementary resources we quantify variation between sexes as well as substantial differences in fine scale patterns among chromosomes and populations.

Results

We performed variant identification on 502 whole genome sequences sequenced in medium to high coverage (mean=16, range=8 to 43). Samples originate from 19 distinct localities spanning the Western Palearctic distribution of the species (Table S2). After filtering we

retained 26,933,469 variants for the whole dataset and used a subset of those for each analysis below (Table S4).

Linkage groups and recombination rate of the barn owl

After aggressively filtering the variant set we ordered 154,706 variants along the 38 largest scaffolds of the genome assembly to create a linkage map for the barn owl. Based on segregation of these markers in 250 individuals from 28 families, we identified 39 linkage groups. The linkage groups identified correspond to scaffolds in the genome assembly for all but two exceptions: Super-Scaffold 2 was split into two linkage groups and Super-Scaffold 3 and 49 were merged in one. The merge of two scaffolds involved one (49) with very few markers and was not resolved (see Supplementary text). The genome assembly of the barn owl therefore contains the sequence of 39 linkage groups out of 45 expected pairs of autosomal chromosomes (Table S1). For these 39 linkage groups the final sex-averaged linkage map spanned 2,066.81 centiMorgans (cM) over a physical sequence of 1,066 million base pairs (Mb) representing 88% of the genome assembly. Therefore the genome average estimate of recombination rate for the barn owl is approximately 1.94 cM/Mb.

The genetic length of linkage groups increases with their physical length (Figure 1A). Each crossover per meiosis adds 50cM to the genetic length of a linkage group. The linkage groups of the barn owl showed a genetic length between 50 and 100 therefore most barn owl linkage groups recombine between once and twice per meiosis. While the slope of the regression of the genetic length on the physical length is significantly positive ($\beta = 0.417, p = 0.004$), the intercept is less than the expected minimum of 50cM under one obligate crossover per chromosome. Several linkage groups between 20 to 40 Mb long have an inferred length of less than 50cM, implying less than a crossing over per meiosis.

Subtle heterochiasmy

To infer heterochiasmy, we look at the sex-specific linkage map estimates (Figure S6 & Table s1). Female barn owls have a 5% larger genetic map than male barn owls with a map length of 2,124cM while the male map is 2,013cM. There appears to be no consistent pattern of heterochiasmy at the linkage group scale (Figure 1B). To investigate potential differences in localisations of crossovers we look at the positioning of inferred crossing over along the length of all chromosomes (Figure 1C). Overall in the barn owl, crossovers tend to occur closer to the linkage group ends than in the middle. However, males tend to recombine more at the extremities of the linkage groups while females more towards the middle.

Because a crossover at the middle of the linkage group will shuffle more markers than a distal crossover, we quantified the rate of intra-chromosomal shuffling (\bar{r}_{intra}) as defined in Veller et al., 2019. Briefly, this quantity measures the relative shuffling of alleles due to a crossover along the length of the chromosome. Thus a crossover in the middle of the chromosome shuffles more alleles than a distal one, and a chromosome with more markers has a higher rate of shuffling than a small one. We estimated rates of intra-chromosomal shuffling in males and females (Fig 1D). Despite recombining more often toward the distal parts of the linkage groups, males show up to 50% higher intra-chromosomal shuffling for larger linkage groups while females show higher rates in intermediate to smaller linkage groups.

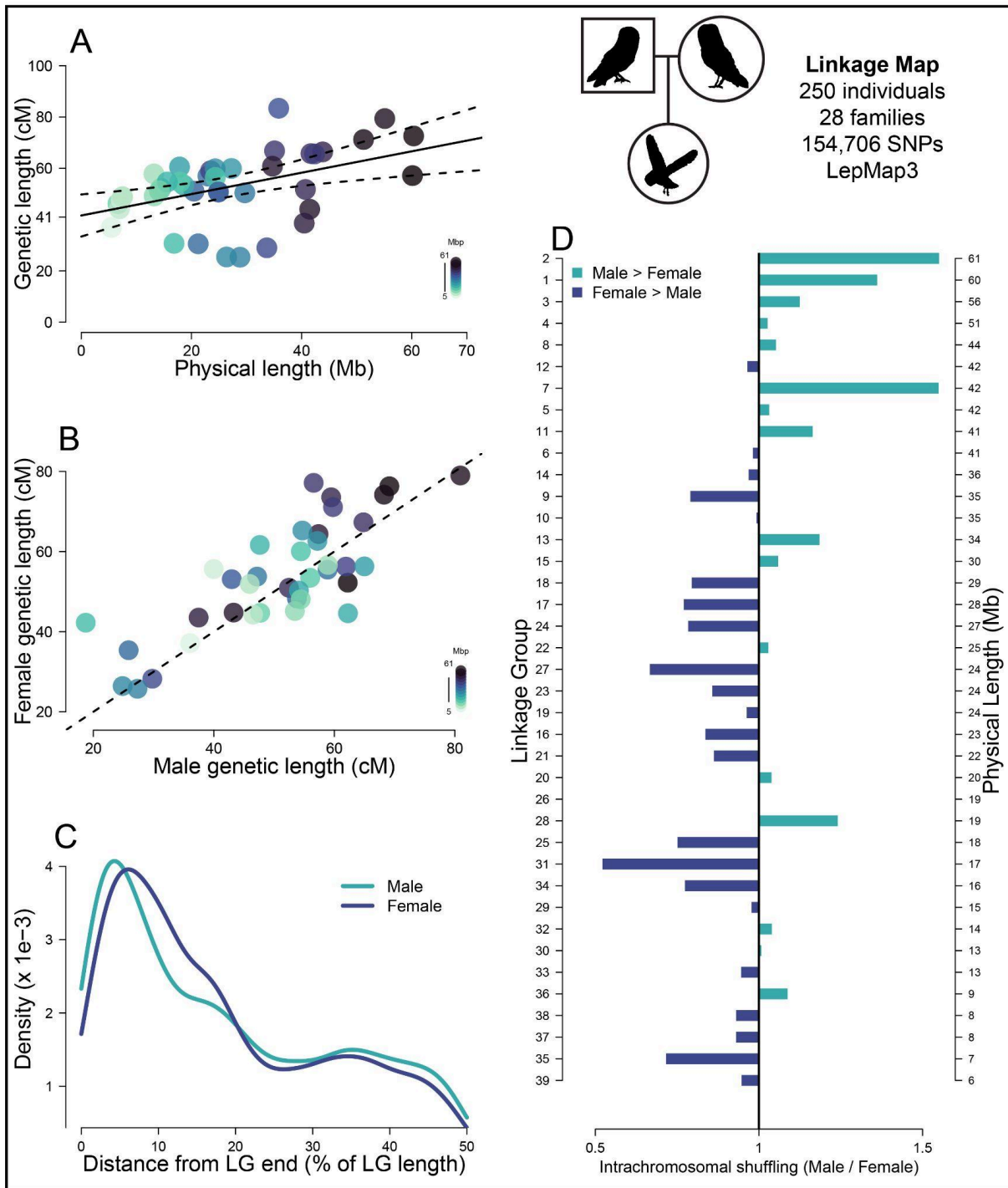


Figure 1 - A linkage map for the barn owl sheds light in fine scale heterochiasmy
 All plots in this figure are products of the linkage map dataset consisting of 250 individuals in 28 families, signified with the pedigree of owls symbol on the top right. **A:** Linkage mapping estimates of sex-averaged genetic lengths for the linkage groups identified in the barn owl assembly plotted against their physical length. Regression line is shown with prediction intervals ($\alpha=41.6$, $\beta=0.4171$, $t=3.093$, $p=0.004$). Colour intensity scales with linkage group physical lengths as in legend. **B:** recombination map length (cM) of linkage groups for females plotted against the recombination map length for males. Dashed line is the identity ($y=x$) line. Each dot represents one linkage group and the colour intensity scales with their physical lengths. **C:** Density plot of male (blue) and female (orange) crossover (CO) counts plotted along the distance from the LGs end. X-axis is in percentage of total linkage group sequence. Density values are scaled so that they sum to 1. **D:** Differences between sexes in rate of intrachromosomal shuffling (r intra) presented as a ratio (male r intra / female r intra) for different LGs. Bars to the right of the black line coloured blue signify higher intra-chromosomal

shuffling in males, while bars to the left of the black line coloured in orange correspond to linkage groups with higher shuffling in females. LGs are ordered by decreasing physical length (larger LG on top) as signified on the second y-axis on the right.

Fine scale variation among linkage groups

To investigate finer scale variation in recombination rates we turn to recombination rates estimated from patterns of linkage disequilibrium (LD) (Figure S5). We estimated recombination rates using ρ in a set of variants identified along the whole genome 76 unrelated birds from Switzerland (CH). The total genetic length estimated from LD was 957 cM, 2.1 times less than the linkage map estimate for the same population. We scaled the total length inferred from LD to equal that of the linkage mapping estimate, to account for the confounding effect of N_e and compared the estimates in non overlapping 1 Mb windows. The correlation at the 1 Mb scale was high ($r = 0.88$, 95% CI: 0.869 - 0.896) with ρ showing higher estimates in regions of low recombination as expected from our limited number of meioses observed (Figure 2A).

The recombination landscape differed among chromosomes with different sizes (three linkage group recombination landscapes highlighting differences in 10kb windows are shown in Figure 2B). To further quantify this variation we looked at the proportion of genomic sequences where recombination occurs. By ordering all 10kb windows for each linkage group based on decreasing recombination rates we calculated the cumulative recombination percentage that occurs in a cumulative percentage of sequence (Figure 2C). Overall, 80% of recombination occurs in approximately 35% of the sequence (dashed grey line in Figure 2C). However, there is substantial variation in the distribution of recombination among linkage groups. To further quantify this skewness we used the Gini coefficient of recombination rates for each chromosome. Briefly, the Gini coefficient corresponds to the area between each curve in Figure 2C and the $y=x$ line with smaller values showing an evenly spread landscape of recombination and higher values a more concentrated one. The genome-wide average Gini coefficient is 0.6 and the linkage group estimates varied between 0.37 and 0.70 (respectively marked with a blue square and a green triangle in Figure 2). Along with a

linkage group of intermediate Gini coefficient (0.57, aquamarine diamond in Figure 2), their landscapes are presented in Figure 2B. We found that the Gini coefficient depends on the physical length of the linkage group with more evenly spread (and elevated) recombination rates in smaller linkage groups and more concentrated landscapes in larger ones (Figure 2D) but the effect seems to diminish as the length increases further. The Gini coefficient also correlates negatively and strongly with the average nucleotide diversity of the linkage group with more concentrated recombination peaks leading to lower average nucleotide diversity (Pearson's $r = -0.9$, 95% CI: -0.947, -0.82) (Figure 2E). Overall, recombination rates vary substantially among the different linkage groups of the barn owl assembly.

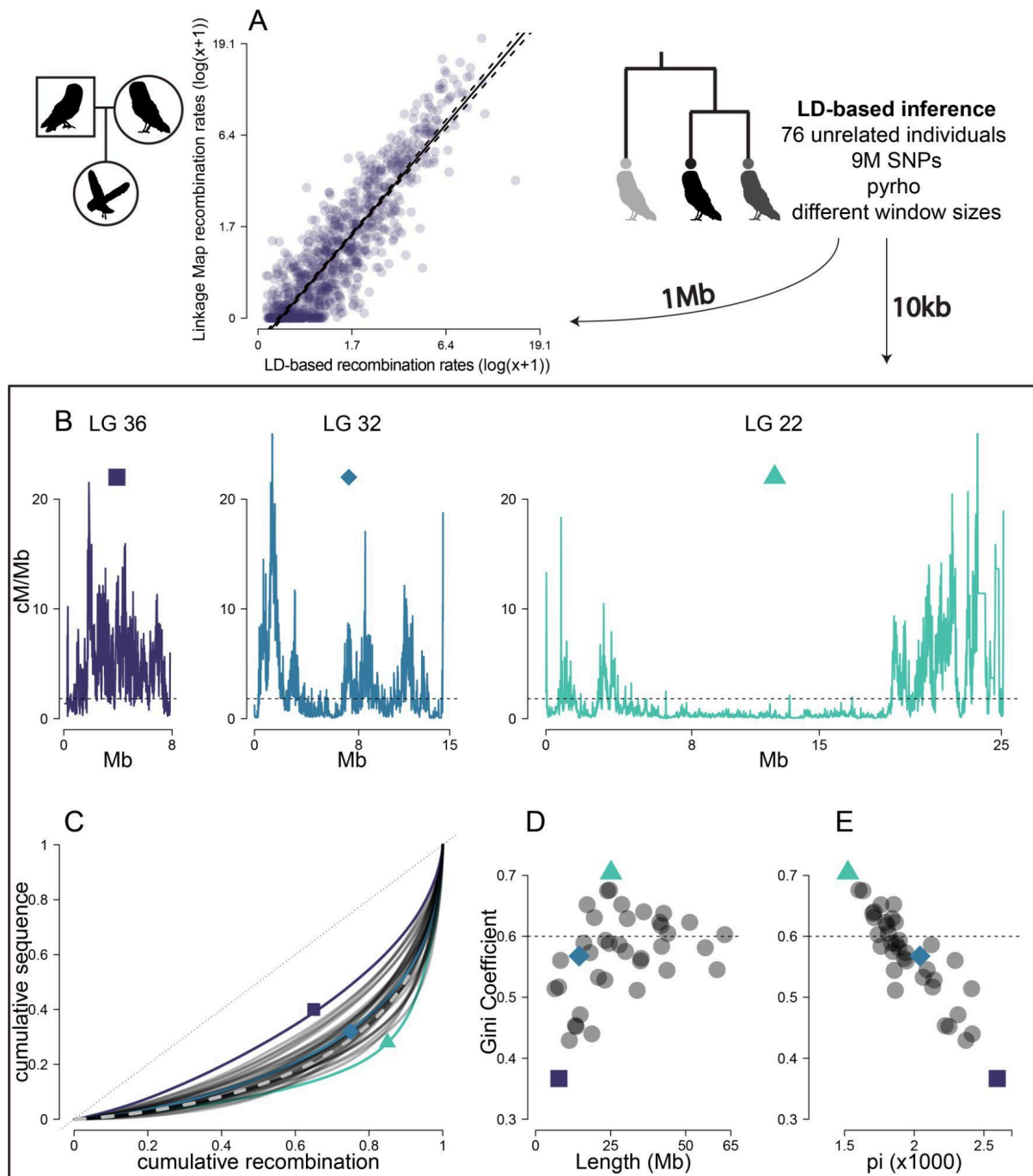


Figure 2 - Variation of recombination among linkage groups.

As in figure 1, pedigree of owls symbol signifies results from linkage mapping. On this figure and later figures the tree of owls shows results from pyrho along with the window resolution used next to it or above it, here 1Mb and 10kb. **A:** Comparison of recombination rate estimates from linkage mapping and LD inference for the Swiss population. The comparison is made in 1Mb windows to avoid inaccurate linkage mapping estimates due to limited meioses observed. Axes are in the natural logarithm of the value + 1 to limit values to 0. Regression line is shown with prediction intervals as dashed lines ($\alpha=-0.285$, $\beta=1.2$, $t=62.92$, $p<0.001$) **B:** The recombination frequency (cM/Mb) in 10kb windows along the physical map of three example linkage groups. The dashed horizontal line is the genome average recombination rate. These linkage groups represent linkage groups with different recombination landscapes. The purple square linkage group has the most equally spread recombination along its length, the blue diamond has an intermediate spread that corresponds to the

genome average and the aquamarine triangle has the most punctuated landscape. **C**: Cumulative sequence plotted against cumulative ordered recombination length for each linkage group (dark grey curves) and genome-wide (grey dashed curve). The black dotted line is the identity ($y=x$) line. The Gini coefficient corresponds to the area delimited by each curve and the identity line. **D**: The Gini coefficient of recombination rates for each linkage group plotted against its physical length. Dashed grey line is the genome average Gini coefficient. **E**: The Gini coefficient of recombination plotted against the average nucleotide diversity of each linkage group. Dashed grey line is the genome average Gini coefficient.

Identifying hotspots of recombination

Because birds lack the PRDM9 gene, recombination hotspots are expected to localise to transcription start and end sites (TSS, TES respectively), as well as CpG islands (CGIs) (Baker et al., 2017; Singhal et al., 2015). To verify this with the barn owl dataset, we used estimates of recombination frequency in non-overlapping windows of 1'000 base pairs (1kb) along the genome. Windows that were annotated to contain either a TSS or a TES ($n=30'224$, 2.7% of windows) or contained a CGI spanning the whole window ($n=13'841$, 1.2% of windows) were identified and their recombination rate was divided by the average in 40kb upstream and 40kb downstream of the focal window (relative recombination rate in 80kb - RRR80). The results showed elevated recombination rates in the focal windows compared to their vicinity (Figure 3A).

The partition of recombination rates along the sequence illustrated in Figure 2D supports the existence of hotspots in the barn owl recombination landscape. We thus looked for recombination hotspots at the kilobase resolution. We define local hotspots as 1kb windows that exhibit 5 times the average recombination rate in 80kb around the focal window ($RRR80 \geq 5$, Figure 3B) and global hotspots as 1kb windows that show a recombination rate higher than 10 times the genome average ($RRR \geq 10$ - Figure 3C). In the Swiss population, we identified a total of 3'949 local hotspots containing 1.8% of the total genetic length and 4'440 global hotspots containing 5.5%. 499 windows were annotated as both local and global hotspots. Local hotspots were usually identified in regions of lower recombination rates which are found towards the middle of linkage groups (Figure 1C) while global hotspots were in peaks of recombination by definition, concentrated around the ends (example in Figure

3D, Figure 3E). Both hotspot classes showed higher GC content distribution compared to the genome average (Figure 3F) which supports their annotation. GC content was higher in global hotspots than local ones.

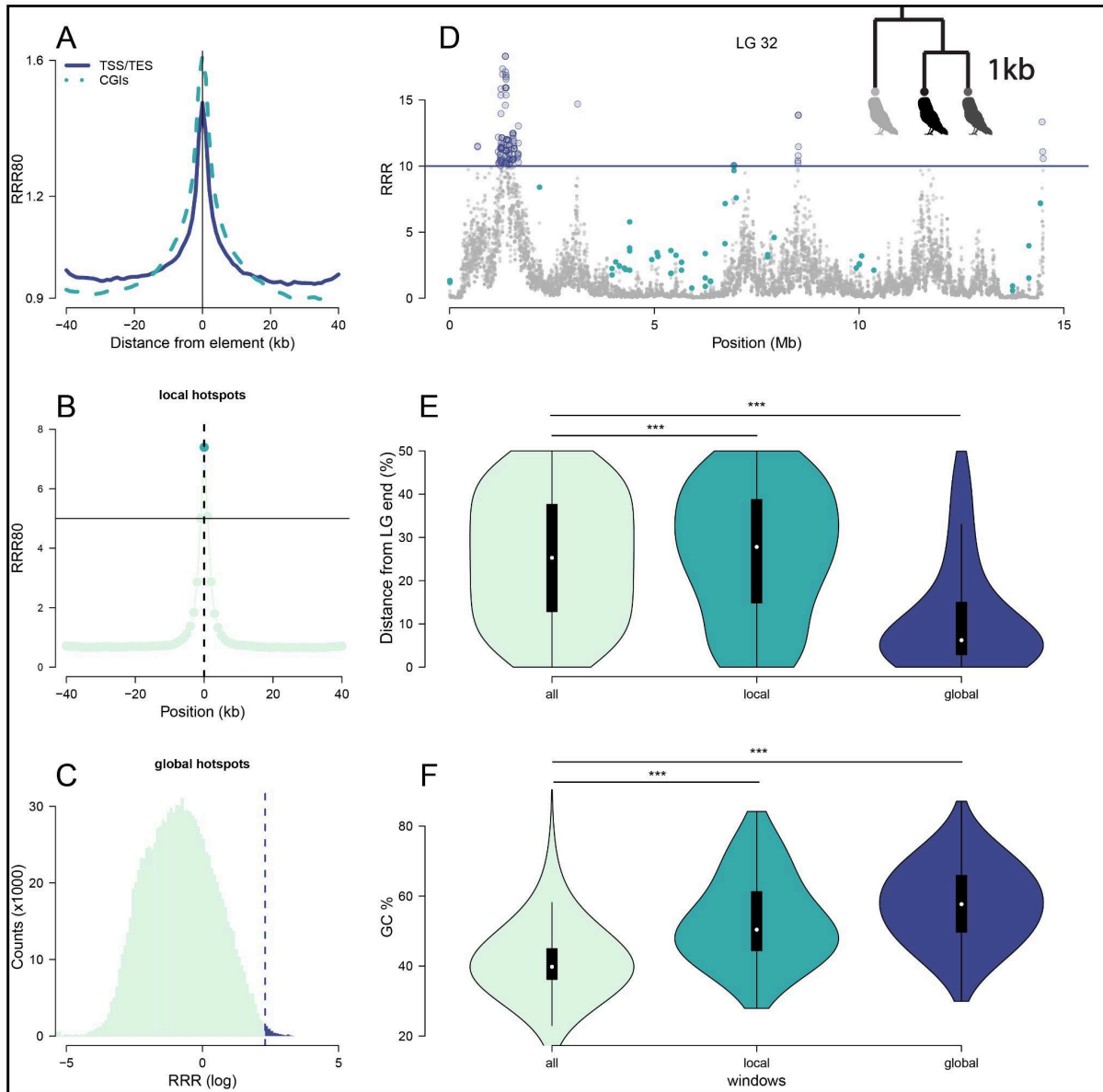


Figure 3. Hotspot characteristics of the barn owl in 1 kb windows.

All results on this figure are using pyrho in 1kb windows as signified with the tree of owls. **A:** Lacking PRDM9, recombination is increased at the local regions around annotated transcription start and end sites (TSS, TES, full line) as well as CpG islands (CGIs, dashed line). RRR80 is the recombination rate of each 1kb window divided by the average in 80kb around. The lines show the average across all identified elements. **B:** Local hotspots are defined as 1kb windows with a recombination rate at least five times higher than the average in 80 kb around ($RRR80 > 5$). This example (which is an average plot of all windows annotated as local hotspots) the focal window (aquamarine dot at position 0) would be annotated as a local hotspot. **C:** Global hotspots are defined as 1kb windows with a recombination rate at least ten times the genome-wide average ($RRR > 10$). All windows right of the dashed line are global hotspots. The RRR is given in log scale for better representation **D:** Example of annotated global and local windows in linkage group (LG) 32. Aquamarine dots are local hotspots

while the purple line shows ten times the global average and thus all windows above the line are global hotspots as signified with the open purple circles. **E**: Violin plot of the location of all windows, local and global hotspots relative to the end of the linkage group. Local hotspots are usually found towards the middle of the linkage groups because they are identified in regions of lower background recombination. On the other hand global hotspots are concentrated around the ends (as expected from Fig 1C). Significance assigned through a Wilcoxon rank sum test. **F**: GC content is elevated in both local and global hotspots. Significance assigned through a Wilcoxon rank sum test.

Low repeatability of fine-scale recombination landscapes

To quantify the change of the recombination landscape in the species we used three populations from the Western Palearctic, Portugal (PT, n=13) and Great Britain (GB, n=13) and Switzerland (CH, n=76) (Figure 4A). Because our sample size in Switzerland was far greater than the other two and to test the replicability of our results, we randomly subsampled two sets of 13 individuals from Switzerland, creating pseudo-replicate populations (CH13, CH13_2). For Portugal and Great Britain genetic length estimates were 1271 and 1345 cM respectively. These estimates matched closely with the ones for the subsampled sets from Switzerland which were 1327 and 1296cM respectively. For all populations, we scaled the results so that the total genetic length would match that of the linkage map estimate of 2067cM. After scaling, we compared all populations with the linkage map in 1Mb windows along the genome and found that genome-average correlations were higher than 0.8 (Figure S1). At finer scale however, recombination landscapes varied among populations (An example landscape for all populations is presented at the 1kb scale for the first five Mb of LG 32 (Figure 4B)), with estimates in Great Britain (GB) showing reduced resolution at the finer scales (blue line in Figure 4B), probably because of reduced genomic diversity due to low sample and historical effective population size (Table S3 & Figure S2, see also Machado, Cumer, et al., 2022).

We quantified the divergence of the recombination landscapes as the correlation of recombination rates on different window sizes (1kb and 100kb) among all pairs of populations (Figure 4C). Two main patterns emerge from the comparison of population landscapes. First, correlation of landscapes depended on the scale used, with smaller windows showing smaller absolute values (above and below diagonal in Figure 4C). Second,

in both scales, correlations of the full Swiss dataset (CH, n=76) with PT and GB were larger than correlation from the subsampled Swiss datasets (CH13, CH13_2). However, as expected the Swiss datasets (CH, CH13, CH13_2) resembled each other more closely (Figure 4C and Figure S3). Concerning hotspots, datasets with smaller sample sizes exhibited more local hotspots than the full Swiss dataset (Figure 4D). However, hotspot overlap was poor between any pair and among all three populations (Figure 4D). This pattern was replicated when comparing hotspot sharing among the Swiss datasets (Figure S4). In fact CH and CH13 shared a lower percentage of local hotspots than did PT and CH. In general, global hotspots showed higher values of sharing than local hotspots and more consistent patterns of sharing.

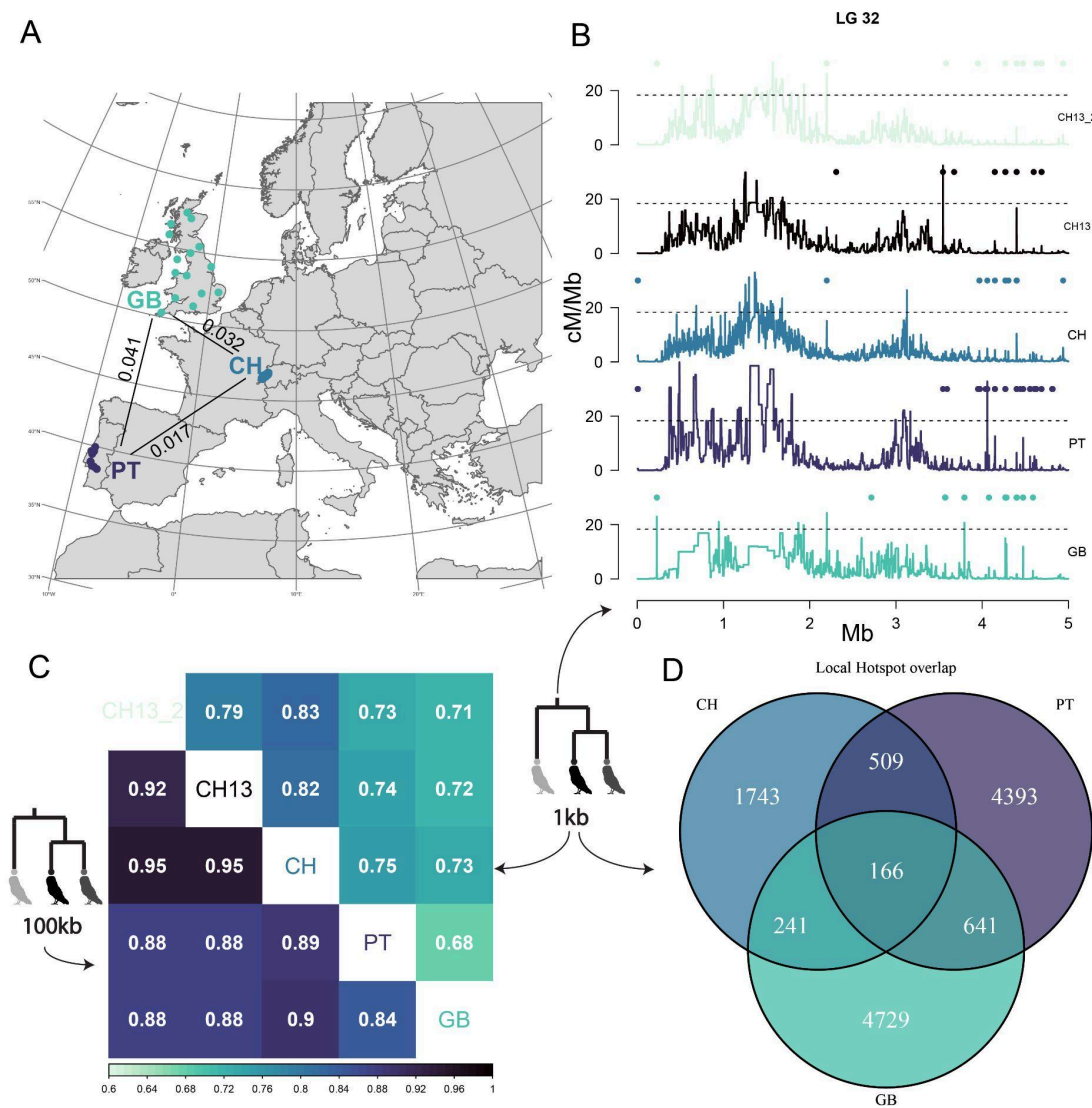


Figure 4. Comparison of recombination landscapes between populations.

A: Map of sampled populations. Numbers correspond to genome-wide pairwise F_{ST} values (Machado, Cumer, et al., 2022) **B:** Example of recombination landscape for 1kb windows in all datasets for the first 5Mb of LG 32. Points above each plot show local hotspots for each population. Local hotspots are identified as in Figure 3 ($RRR80 >5$). The dashed line is the line above which we defined global hotspots (10 times the average for each population). **C:** Correlation matrix of recombination rates for 1 kb windows (above diagonal) or 100 kb windows (below diagonal). **D:** Venn diagram of local hotspot counts within and between populations. Population codes are as follows: CH: full Swiss dataset (n=76), PT: Portuguese dataset (n=13), GB: Great Britain dataset (n=13), CH13: undersampled first Swiss dataset (n=13), CH13_2: undersampled second Swiss dataset (n=13).

Discussion

Recombination is a major mediator of evolution but we know little about the recombination landscape of most species. In this study utilising an extensive whole-genome-sequencing (WGS) dataset of a pedigreed population we apply two recombination inference methods, linkage mapping and an LD-based approach to describe broad and fine scales of recombination variation. We show how in the barn owl, heterochiasmy exists not on the broad scale but rather in subtle, fine-scale differences in crossover placement and shuffling proportions. Furthermore, despite few (1-2) crossovers per chromosome, we find large variation in recombination rates among linkage groups that shapes the diversity landscape and is only partially determined by their physical length. We show that on the kilobase scale, recombination is increased in windows that contain transcription start and end sites and CpG islands as expected from a species without PRMD9 (Baker et al., 2017). At the same time, local hotspots are found in regions of lower average recombination (usually in the middle of linkage groups) and exhibit lower GC ratio compared to global hotspots. Lastly, population comparisons show sharing of only a few local hotspots despite low genetic differentiation and high broad-scale recombination correlation. We discuss these results and their implications below.

Linkage groups in the barn owl assembly

Complete genome assemblies are a precious resource that requires multiple sources of information. Beyond read acquisition, assembling a genome requires an ordering process to

orient the reads into scaffolds and the scaffolds into chromosomes. In-silico this can be achieved using physical mapping of the reads (Burton et al., 2013; Lieberman-Aiden et al., 2009), or through the use of linkage mapping (Fierst, 2015) although linking computationally assembled scaffolds to karyotypic chromosomes will eventually require molecular techniques such as FISH (Shakoori, 2017). The latest barn owl assembly was assembled into super-scaffolds using optical genome mapping (BioNano) (Machado, Cumer, et al., 2022). In this study we verified and improved the barn owl assembly by anchoring the largest 38 scaffolds into 39 linkage groups and revealed that the genome assembly of the barn owl is of chromosome-level quality.

The karyotype of the barn owl contains 45 autosomal pairs (Belterman & De Boer, 1984; Peona et al., 2018). This implies that the linkage map is still missing 6 autosomes. We note that these elements might be partially present in the physical assembly since smaller scaffolds with a few tens of identified markers that passed filtering could not be confidently allocated to linkage groups. Regardless, the chromosomes missing are probably the smallest six microchromosomes, or dot chromosomes, notoriously difficult to sequence and assemble due to high GC content and reduced chromatin accessibility (Bravo et al., 2021; Burt, 2002; Waters et al., 2021). Notably, such elements were only recently assembled in the chicken genome (Huang et al., 2023) and are missing from most available bird reference genomes (Peona et al., 2018). These elements will remain elusive until future studies make use of advances in long-read technologies (Marx, 2023) to complete the reference genomes of birds. In this endeavour, linkage mapping, when available, can be an invaluable tool.

Heterochiasmy beyond the broad strokes

Our results point at a fine scale variation in crossover placement between sexes that is masked by broad scale differences. Consequently, this different placement of crossovers between sexes leads to a differential shuffling of markers between sexes and among chromosomes. Most avian studies of recombination to date have examined total genetic

lengths in males and females and their results are inconclusive on a general pattern of heterochiasmy in the class. For example male collared flycatchers exhibit higher genetic lengths than females while in a recent study of the great reed warbler (*Acrocephalus arundinaceus*) no large differences are found between sexes (Kawakami et al., 2014; Zhang et al., 2023). On the other hand sparrows and great tits show a female dominated recombination landscape (McAuley et al., 2023; van Oers et al., 2014). Only recently has there been an appreciation of the fine-scale variation between sexes in studies of the great reed warbler and the sparrow (McAuley et al., 2023; Zhang et al., 2023). This is an important shift since such fine scale variation can impact effective genetic shuffling between the sexes with implications for adaptive potential (Brekke et al., 2023; McAuley et al., 2023; Veller et al., 2019). Although fine-scale information is not available for other bird species making a comparison impossible at the moment, an emerging pattern is that broad scale descriptions might not reveal the whole picture of heterochiasmy in birds and more thorough quantification of sex-specific recombination is required.

Identifying the causes of such heterochiasmy is a complicated task and multiple factors can affect recombination in each sex (Sardell & Kirkpatrick, 2019). Beyond mechanistic processes (Brick et al., 2018; Kong et al., 2004; Phillips et al., 2015; Tease & Hultén, 2004) and standing genetic variation (Halldorsson et al., 2019; Johnston et al., 2016; Kong et al., 2008) heterochiasmy can also be adaptive. Of the hypotheses put forward, the meiotic drive hypothesis (Brandvain & Coop, 2012) which examines female-specific evolution of recombination rates to counteract the effect of meiotic drive, seems to provide a reasonable explanation for the most typical pattern observed, the increased crossover frequency of females closer to the centromeres and males closer to telomeres. The theory has also received empirical support in some species (Johnston et al., 2017). In the barn owl, males recombine more towards the edges of the linkage groups but all autosomal chromosomes in the species are acrocentric or telocentric (Belterman & De Boer, 1984). Therefore we can draw no conclusion on this pattern unless we know which end corresponds to the centromere and which to the telomere. To our knowledge, in the avian clade, heterochiasmy

has yet to be associated with centromeres and telomeres although a male-biased recombination at ends of linkage groups has been identified (this study and Zhang et al., 2023). This lack of conclusive results can be linked with a poor genomic annotation of centromeric and telomeric sequences in the genome assemblies. Future work should concentrate on properly annotating available resources and looking for the possible causes of fine-scale differences between sexes.

Variation among linkage groups

Our results show that barn owl linkage groups recombine at most twice per meiosis. This result is in line with an expectation of one cross-over per chromosome (or chromosome arm) and the generally small acrocentric (or telocentric) chromosomes in the barn owl karyotype (<70Mb) (Coop & Przeworski, 2007). This finding contrasts with results for some bird species. The biggest chromosomes of the chicken and the flycatcher approach 300cM in genetic length, while studies in other passerines agree more with estimates of the barn owl (Groenen et al., 2009; Kawakami et al., 2014). The source of this variation in the order is unknown. Reasonable hypotheses include the localised suppression of recombination in some species (for example through segregating structural variations like inversions) or inter-specific variation in the strength of crossover interference (Kirkpatrick, 2010; Otto & Payseur, 2019).

Some linkage groups showed evidence of less than one crossover per meiosis. On the one hand this can be a true signal facilitated through less than one crossover per bivalent. The absence of the obligate crossover can lead to aneuploidy which coupled with the linkage groups' intermediate size can probably generate severely deleterious consequences. On the other hand this observation can be due to chromosome parts missing in the assembly or filtered out during quality control, which can lead to missed distal crossovers. A larger sample size and/or a more complete assembly that incorporates the distal parts of all chromosomes might help identify the cause.

In our study, recombination rates vary substantially between and within chromosomes. As expected from the obligate cross-over mentioned above, smaller chromosomes tend to have higher rates of recombination compared to longer chromosomes. Furthermore, longer chromosomes show a more pronounced variation in recombination rates along their length. Specifically, they tend to follow a U-shaped pattern, with reduced recombination in their middle, regardless of centromere position. In our results, this effect seems to diminish with increasing length implying that in the barn owl, even longer chromosomes would not further impact the skewness of recombination rates. Recently a hypothesis has been put forward to explain the dependence of recombination rate skewness and chromosome length, based on the position of telomeres, centromeres and one crossover per chromosome (Brazier & Glémin, 2022; Haenel et al., 2018). It is still unknown however, if this distinction between high and low recombination regions follows a compartmentalisation of the genomic sequence into active and inactive chromatin and how this broad scale pattern defines or is guided by a fine-scale hotspot landscape (Hildebrand & Dekker, 2020; Jerkovic´ & Cavalli, 2021). Further, it is also unknown to what extent such variation extends to other species.

Even if the causes of such patterns within and between chromosomes remain a mystery some of their consequences can still be glimpsed. The most striking consequence of the unequal distribution of recombination rates along a specific length of sequence is the impact it has on nucleotide diversity. In the linkage groups studied, the Gini coefficient correlates very well with the average nucleotide diversity. Such an outcome is expected through the action of linked selection (Begun & Aquadro, 1992; Charlesworth & Jensen, 2021). If recombination is spread throughout the length of the sequence, alleles are uncoupled faster from selected variants allowing an increase of standing variation. On the contrary, long stretches of reduced recombination, through the action of linked selection lead to reduced diversity (Charlesworth et al., 1993; Charlesworth & Jensen, 2021). This reduced diversity can have multiple implications. It can impact homozygosity and lead to extended runs of homozygosity (ROH), as seen in other studies through a positive correlation of ROH and recombination rates (Hewett et al., 2023; T. J. Pemberton et al., 2012). Reduced diversity

can also affect estimates of divergence between populations, measures often used to identify local adaptation leading to biases and misleading inference (Booker et al., 2020; Burri, 2017; Charlesworth, 1998). Therefore, such variation within and between chromosomes should be accounted for, when attempting to infer processes that act on the sequence. Nevertheless, tackling these issues requires a careful examination of resources. For example, an LD-based map like the one we use in this study will always be influenced by processes that impact LD like selection (Kim & Nielsen, 2004; O'Reilly et al., 2008). Therefore, its use to correct for recombination when inferring selection will introduce biases. On this end, other methodologies like the linkage map approach or sperm typing might prove more useful despite being less broadly applicable (Peñalba & Wolf, 2020).

Hotspots & population comparison

A major distinction among the recombination landscapes of species studied thus far is the presence or absence of the PRDM9 gene. PRDM9 directs the recombination machinery in specific genomic regions through the generation of H3 lysine K3 tri-methylation marks (H3K4me3). These marks attract the SPO11 protein which initiates the formation of double strand breaks, the precursor to homologous recombination (Arter & Keeney, 2023; Lam & Keeney, 2014). In species that lack the PRDM9 gene, including all birds, H3K4me3 marks are concentrated in regions of accessible chromatin like promoter regions of genes (Baker et al., 2017). Because these regions have an increased frequency of recombination and co-localise with CpG islands (CGIs) and transcription start or end sites (TSS, TES), a correlation of these genomic elements with recombination is often observed in species without PRDM9 (Auton et al., 2013; Baker et al., 2017; Kawakami et al., 2017; Lam & Keeney, 2015; Schield et al., 2020; Singhal et al., 2015). Our results illustrate a marked increase around these regions, confirming the hypothesis that recombination is concentrated in their close proximity.

While recombination hotspots have been identified through the action of PRDM9, species in which the gene is absent do not necessarily harbour hotspots (Kaur & Rockman, 2014; Smukowski Heil et al., 2015). Usually the presence or absence of hotspots is inferred from the skewness of recombination rates along the genome (Myers et al., 2005). For example in humans a Gini coefficient of 0.8 shows a marked discrepancy between the hotspot and non-hotspot regions of the genome (Kong et al., 2010; Myers et al., 2005). In *Caenorhabditis elegans* the Gini coefficient is 0.278 and recombination is spread along the full genomic sequence (Kaur & Rockman, 2014). Barn owl chromosomes harbour values across the spectrum (between 0.36 and 0.7). However, a genome-wide Intermediate value inferred here (0.6), is obviously harder to place in one or the other category so to verify the presence or absence of hotspots in our dataset we turned to identifying them. However, recombination hotspots are not clearly defined in the literature. Thus, we used two definitions for a hotspot, local and global hotspots. Local hotspots are defined based on a relative increase (ex. 5x, 10x) of recombination rates compared to the surrounding region (Myers et al., 2005; Singhal et al., 2015). We indeed identified local hotspots but they were mostly located in regions of low recombination. Past simulation work of power of hotspot inference shows that this is expected to be the case because power diminishes as recombination increases (Singhal et al., 2015). On the other hand global hotspots defined following Halldorsson et al., 2019 were found in different genomic regions than local ones. Both classes were supported by an increase of GC content, either through the action of GC-biased gene conversion or through their co-localisation with GC-rich regions as mentioned before (Eyre-Walker, 1993).

A last implication of missing the PRDM9 gene is the evolutionary stability of recombination hotspots. The Zinc-finger domain of PRDM9 evolves quickly changing its target sequence leading to differences in the localisation of hotspots between species and individuals (Axelsson et al., 2012; Kong et al., 2010; Myers et al., 2010). On the other hand in its absence the functional regions that 'attract' the recombination machinery remain stable for sometimes millions of years leading to a stable fine-scale recombination landscape (Lam & Keeney, 2015; Singhal et al., 2015). The local hotspots inferred in this study showed very

small overlap between pairs of populations. Even when comparing subsets of individuals from the same population (CH with CH13 and CH13_2) hotspot sharing was lower than 50%. This observation generates doubt about the biological usefulness of local hotspots in this study setting. While some local and global hotspots coincide, and the GC content increase provides support for the existence of some local hotspots, the set identified is expected to harbour multiple false positives. Furthermore, since the number of hotspots identified was higher in datasets with a smaller sample size, an effect of sampled genetic diversity and inference noise cannot be ruled out.

Beyond hotspot sharing, the similarity of recombination landscapes was only validated at broad scales. Our studied populations diverged after the last glacial maximum when they expanded out of an iberian refugium (Cumer, Machado, Dumont, et al., 2022a; Machado, Cumer, et al., 2022). This timescale coupled with the intermediate dispersal abilities of the species has led to a shallow genetic differentiation (Altwegg et al., 2003; Machado, Cumer, et al., 2022). This lack of a convergent fine-scale recombination landscape is not expected from a species without PRDM9. However, we are cautious in interpreting such results as a true divergence of the fine-scale recombination landscape. The dependence of inference on the LD patterns and standing variation can confound results, especially in the finer scale where statistical noise increases. Thus, whether this result supports a divergent fine-scale landscape or method limitations remains unclear. Future work should be cautious when using fine scale estimates in non-model species and might benefit from verification of either the recombination itself or at least the noise around inference as illustrated here with the use of subsampled datasets.

Conclusion

In this study we elucidated the recombination landscape of the barn owl genome. We used two complementary approaches, linkage mapping and LD-based inference to overcome the limitations of each method. We showed that sex-differences in recombination exist but only beyond the broad scale and that variation among linkage groups is a significant determinant

of diversity. Further, while we show the existence of hotspots in a species without PRDM9 we caution the use of fine-scale inferences and the expectation of the conservation of a fine-scale recombination landscape between populations.

Methods

Samples and Sequencing

A total of 333 barn owl samples from Switzerland were sequenced for this study. In Eastern Switzerland breeding of barn owls has been monitored for over 30 years with the installation of nest boxes. During the breeding season, the nest boxes are controlled for occupancy every 4 weeks. If a nest is discovered, a second visit is planned after all eggs have hatched and a third visit at 55 days, right before the first chick finishes fledging and flies away from the nest. Individuals are ringed and blood is extracted from their brachial vein. Adult parents are also captured when possible and subjected to the same treatment. Using the ring identifiers of parents and offspring and the fact that barn owls show rare extra-pair paternity (Roulin et al., 2004) an observational pedigree has been constructed for the population.

Sequencing was spread throughout 2020 and 2021. 285 individuals belonging to families based on pedigree information were sequenced in 2020 and 2021. Initially, to construct a linkage map we attempted to sequence families that had many offspring and grandparent information. Sample DNA was extracted from blood using DNeasyBlood & Tissue kit (Qiagen) following manufacturer's instructions. Libraries were sequenced with Illumina HiSeq 4000. We increased the dataset of sequenced individuals by choosing 48 more samples of owls that had the maximum number of descendants based on the field pedigree. We sequenced these samples in 2021. DNA extraction was performed with DNeasyBlood & Tissue kit (Qiagen) and sequencing was performed using Illumina NovaSeq 6000. All sequencing took place at the Lausanne Genomic Technologies Facility (GTF, University of Lausanne, Switzerland).

Variant Discovery & Filtering

All available barn owl sequences were used for variant discovery. This included individuals mentioned above and samples from previous sequencing efforts (Cumer, Machado, Dumont, et al., 2022a; Cumer, Machado, Siverio, et al., 2022; Machado, Cumer, et al., 2022; Machado, Topaloudis, et al., 2022) along with 6 samples from Georgia and 3 from the island of Corsica (Table S2). In total 502 samples were processed through the variant discovery pipeline described below. Raw reads were processed with trimmomatic v0.39 (Bolger et al., 2014). Sequence adapters were removed and reads with a length less than 70bp were excluded. Mapping was performed with BWA-MEM v0.7.17 (H. Li, 2013) on the barn owl genome assembly (<https://www.ncbi.nlm.nih.gov/nuccore/JAEUGV000000000>) (Machado, Cumer, et al., 2022)) and read groups were added with samtools v1.15.1 (H. Li et al., 2009). Since the GATK v4.2.6 (Auwera et al., 2013) pipeline was used for variant discovery, base quality score recalibration (BQSR) was performed using a truth set previously published in (Cumer, Machado, Dumont, et al., 2022b). GATK's Haplotype caller was run with default parameters for each individual separately to generate individual *gvcf* files.

These files were merged and joint calling was performed on all individuals together using *GenotypeGVCFs*. We initially identified 30,620,917 variants in the dataset. Filtering focused on bi-allelic SNPs and consisted of the core technical filters suggested in the GATK pipeline, a "mappability" mask and a manual individual depth filtering. Specifically, technical filters included the following criteria: $QD < 2.0$, $QUAL < 30$, $SOR > 3.0$, $FS > 60.0$, $MQ < 40.0$, $MQRankSum < -12.5$ and $ReadPosRankSum < -8.0$. A further filtering was the exclusion of regions of the genome where our ability to confidently map reads is limited (i.e. a "mappability" mask) (Corval et al., 2023). Briefly, the reference genome was split into reads of 150 base pairs (bp) with a sliding of 1bp. These artificial reads were mapped back to the reference using *bwa-mem*. Regions of the sequence where less than 90% of the reads did not map perfectly and uniquely were discarded by excluding variants using a bed file. Variants were also filtered based on individual depth. A minimum and a maximum cutoff

were applied. For the minimum cutoff, any genotype with less than five reads supporting it was set to missing (Benjelloun et al., 2019). For the maximum, a distribution of autosomal read depth per individual was extracted for a region (Super-Scaffold_1 and Super-Scaffold_2) with a length of 133.5Mb. The mean and standard deviation of depth was estimated and any genotype with a read depth of more than three standard deviations from the mean was set to missing to avoid the effect of repeated regions. After filtering 26,933,469 variants were kept.

Pedigree and relatedness

The pedigree from observational data was confirmed with genomic information from a subset of the genome. SNPs from three scaffolds (Super-Scaffold_11,12, and 14) were filtered for minor allele count (>5), missing data (<10%) and were pruned for linkage disequilibrium using plink v.1.9 ([Chang et al., 2015](#)) with the command `--indep-pairwise 100 10 0.1`. This filtering created a dataset with 91,874 SNPs. A genomic kinship matrix was calculated using the [Weir & Goudet, 2017](#) method as implemented in hierfstat ([Goudet, 2005](#)) R package. The kinship from genomic data was compared with the pedigree kinship, calculated using the kinship2 ([Sinnwell et al., 2014](#)) R package, and the pedigree was completed by manually resolving the first and second degree links when those could be resolved. Both k1 and k2 statistics in SNPRelate were used to discern between relationships with the same kinship value (ex. parent-offspring and siblings). A set of unrelated individuals was selected automatically by pruning the genomic kinship table to only include individuals with a kinship of less than 0.03125. This method left a subset of 187 unrelated individuals of which 76 were from the Swiss population.

Linkage Mapping

Lep-MAP3 (LM3) ([Rastas, 2017](#)) was used to create a linkage map. A dataset of 250 individuals in 28 families was included where a family in LM3 is defined as a set of

individuals around a unique mating pair. Most families had 4 offspring and 4 grandparents but numbers differed and ranged from 2 to 8 offspring and from 0 to 4 grandparents. To run LM3 a stringently filtered dataset of bi-allelic SNPs was used. Specifically, we removed mendelian incompatibilities using bcftools' mendelian plugin (Danecek et al., 2021) and retained a minimum of 5% MAF and a maximum of 5% missing data. We also filtered out SNPs that were less than one thousand base pairs (1kb) apart using VCFtools (Danecek et al., 2011). The first step of the LM3 pipeline *ParentCall* was used to transform the data into the appropriate LM3 format and the options *halfSibs* and *removeNonInformative* were included. Data was filtered in LM3 using the *Filtering2* command, to shrink the size of the dataset. Specifically *dataTolerance* was set to 0.01 as suggested by the author and *missingLimit* and *familyInformativeLimit* were set to 28. This meant that only variants that were non-missing and informative in all families were kept. After filtering the dataset, we retained 163,950 variants. We used *SeparateChromosomes* to identify the putative linkage groups (LGs) based on a user-defined logarithm of odds (LOD) score cutoff. We selected a LOD score of 15 (for a justification see Supplementary text - making a linkage map). Finally *OrderMarkers* with the *usePhysical* option was executed. Ordering was repeated three times and the output with the best Likelihood was selected for each linkage group. All three runs were compared to test for variation in estimated genetic maps. We also tested the effect of three mapping functions (Morgan's, Haldane's and Kosambi's) on the estimated genetic maps (Figure S9).

In linkage mapping certain markers might be erroneously mapped especially at the extremities of the LGs. Thus all markers with a cM jump higher than 2 in a region of 100 markers around the ends of the LGs were filtered out. A homemade script inspired by *LepWrap* (Dimens, 2022) was used. We also pruned the resulting Marey maps (plot of cM position on physical position) by regressing the genetic order of markers with the physical position order across a linkage group. Markers with an absolute residual value of more than 100 were removed to reduce noise in the resulting maps. Finally we fitted a generalised additive model using the R Package *mgcv* (Wood, 2011) and *scam* (Pya & Wood, 2015) forcing a monotonically increasing smoothing spline. This makes sure that the next cM position will be bigger than the previous one and gives a better fit to the data (Figure S10).

Linkage disequilibrium recombination

To execute SMC++ (Terhorst et al., 2017) we followed the authors' instructions as presented in the software's GitHub page (<https://github.com/popgenmethods/smcpp>). In summary, missing data was re-coded using Plink2 (Chang et al., 2015) and the 5 samples with the highest coverage were selected as distinguished individuals to be provided to SMC++. The command `vcf2smc` was run for each distinguished individual. When executing ``vcf2smc`` the mappability mask was excluded by using the `-m` option. The model was estimated using all output files from the previous step and with a mutation rate of $4.6e-9$ estimated from family data of a collared flycatcher (Smeds et al., 2016). The csv-formatted estimate of piecewise-constant effective population size in past generation intervals was used in subsequent `pyrho` (Spence & Song, 2019) analyses.

We ran `pyrho` with an unphased set of markers for 76 Swiss individuals. The first step in the `pyrho` implementation was the pre-calculation of a two-locus likelihood look-up table. This step takes into account the N_e estimates from `SMC++` (Figure S2). For our Swiss samples our number of diploid individuals was 76 and we used the Moran approximation with a size of 200. After the inference of the lookup table the "hyperparameter" command was run to estimate metrics on the performance of different window sizes and block penalties. The authors' guidelines were followed on how to select the best combination of parameters. Briefly we summed the Pearson correlation statistics outputted by `pyrho` and plotted their total sum against the L2 values. The authors suggest (<https://github.com/popgenmethods/pyrho#hyperparam>) that depending on the implementation one might opt to choose the parameter combination that maximises the correlation measures or minimises L2. In our case both conditions were satisfied with one combination of parameters and we run `pyrho` with that set of parameters. A table of the hyperparameter values for all populations can be found in Table S3. With the inferred hyperparameters, recombination rate was estimated using the `optimise` command on `vcfs` containing individual scaffolds which were previously filtered for singletons and a minimum distance of 10bp between variants as in Wall et al., 2022.

Downstream analyses

Rate of intra-chromosomal shuffling was calculated from recombination rates inferred from mapping distances following Veller et al., 2019. Recombination rate estimates from `pyrho` were averaged across non-overlapping windows of different lengths using a custom script. Windows of sizes 1kb, 10kb, 100kb and 1Mb were created from the reference sequence using `bedtools makewindows` from `bedtools v2.3` ([Quinlan, 2014](#)). These windows were overlapped with the `pyrho` windows and the recombination rate in cM was calculated by multiplying the recombination probability estimate with the length of each interval and then translating this to cM using Haldane's function. For each window, nucleotide diversity was

calculated using VCFtools and the `--window-pi` command. Estimates were corrected for masked nucleotides in each window. Sequence GC content was calculated using the reference sequence and the `bedtools nuc` command. We annotated CpG islands using the UCSC genome browser CpG island annotation tool `cpg_hl` (Kent et al., 2002) with default parameters. The Gini coefficient was calculated using Desctools v.0.99 (Signorell, 2023). Transcription start and end sites were annotated using the genome annotation from [NCBI](#) as the first and last positions of the genomic sequence for each gene. Intersection of different bed files was performed using `bedtools`. Local hotspots were annotated by dividing the estimate of recombination rate in each focal window with the average recombination in 80kb around (40 kb upstream and 40 kb downstream). Global hotspots were annotated as windows with at least 10 times the genome average recombination rate.

Colour palette used is 'mako' from the R-package viridis v.0.6.4 and is consistent throughout the figures (Garnier et al., 2023). Images of owls come from PhyloPic (<https://www.phylopic.org>). Map in Figure 4A was made with tmap v3.3-4 (Tennekes, 2018) using the Natural Earth high resolution dataset. Corrplot v.0.92 was used for correlation plot in Figure 4C (Wei & Simko, 2021). Vioplot v0.4.0 was used to create the violin plot in Figure 4D (Adler et al., 2022). Tidyverse v.2 was used for data management (Wickham et al., 2019). All analyses were executed in R v.4.3.1 ([R Core Team, 2023](#)) using the Rstudio IDE ([Posit team, 2022](#)). Light figure modification was performed in Adobe Illustrator. Scripts with commands used for data generation and downstream analyses can be found in https://github.com/topalw/Chapter_A.

Author contributions

A.T. and J.G. devised the project. A.R. provided the samples. A-L.R. and C.S. carried out the DNA-extraction and library preparation. A.T., E.L., and T.C. generated and filtered the variant dataset. A.T. performed analyses and wrote the manuscript with input from all authors.

Acknowledgements

References

- Adler, D., Kelly, S. T., Elliott, T., & Adamson, J. (2022). *vioplot: Violin plot*.
<https://github.com/TomKellyGenetics/vioplot>
- Altwegg, R., Roulin, A., Kestenholz, M., & Jenni, L. (2003). Variation and Covariation in Survival, Dispersal, and Population Size in Barn Owls *Tyto alba*. *Journal of Animal Ecology*, *72*(3), 391–399.
- Arbeithuber, B., Betancourt, A. J., Ebner, T., & Tiemann-Boege, I. (2015). Crossovers are associated with mutation and biased gene conversion at recombination hotspots. *Proceedings of the National Academy of Sciences*, *112*(7), 2109–2114.
<https://doi.org/10.1073/pnas.1416622112>
- Arter, M., & Keeney, S. (2023). Divergence and conservation of the meiotic recombination machinery. *Nature Reviews Genetics*, 1–17.
<https://doi.org/10.1038/s41576-023-00669-8>
- Auton, A., Li, Y. R., Kidd, J., Oliveira, K., Nadel, J., Holloway, J. K., Hayward, J. J., Cohen, P. E., Grealley, J. M., Wang, J., Bustamante, C. D., & Boyko, A. R. (2013). Genetic Recombination Is Targeted towards Gene Promoter Regions in Dogs. *PLOS Genetics*, *9*(12), e1003984. <https://doi.org/10.1371/journal.pgen.1003984>
- Auton, A., & McVean, G. (2007). Recombination rate estimation in the presence of hotspots. *Genome Research*, *17*(8), 1219–1227. <https://doi.org/10.1101/gr.6386707>
- Auwers, G. A., Carneiro, M. O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K. V., Altshuler, D., Gabriel, S., & DePristo, M. A. (2013). From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. *Current Protocols in Bioinformatics*, *43*(1). <https://doi.org/10.1002/0471250953.bi1110s43>
- Axelsson, E., Webster, M. T., Ratnakumar, A., Consortium, T. L., Ponting, C. P., &

- Lindblad-Toh, K. (2012). Death of PRDM9 coincides with stabilization of the recombination landscape in the dog genome. *Genome Research*, 22(1), 51–63. <https://doi.org/10.1101/gr.124123.111>
- Backström, N., Forstmeier, W., Schielzeth, H., Mellenius, H., Nam, K., Bolund, E., Webster, M. T., Öst, T., Schneider, M., Kempnaers, B., & Ellegren, H. (2010). The recombination landscape of the zebra finch *Taeniopygia guttata* genome. *Genome Research*, 20(4), 485–495. <https://doi.org/10.1101/gr.101410.109>
- Baker, Z., Schumer, M., Haba, Y., Bashkirova, L., Holland, C., Rosenthal, G. G., & Przeworski, M. (2017). Repeated losses of PRDM9-directed recombination despite the conservation of PRDM9 across vertebrates. *eLife*, 6, e24133. <https://doi.org/10.7554/eLife.24133>
- Barton, N. H., & Charlesworth, B. (1998). Why Sex and Recombination? *Science*, 281(5385), 1986–1990. <https://doi.org/10.1126/science.281.5385.1986>
- Begun, D. J., & Aquadro, C. F. (1992). Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. *Nature*, 356(6369), Article 6369. <https://doi.org/10.1038/356519a0>
- Belterman, R., & De Boer, L. (1984). A karyological study of 55 species of birds, including karyotypes of 39 species new to cytology. *Genetica*, 65(1), 39–82.
- Benjelloun, B., Boyer, F., Streeter, I., Zamani, W., Engelen, S., Alberti, A., Alberto, F. J., BenBati, M., Ibnelbachyr, M., Chentouf, M., Bechchari, A., Rezaei, H. R., Naderi, S., Stella, A., Chikhi, A., Clarke, L., Kijas, J., Flicek, P., Taberlet, P., & Pompanon, F. (2019). An evaluation of sequencing coverage and genotyping strategies to assess neutral and adaptive diversity. *Molecular Ecology Resources*, 19(6), 1497–1515. <https://doi.org/10.1111/1755-0998.13070>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Booker, T. R., Ness, R. W., & Keightley, P. D. (2017). The Recombination Landscape in Wild

- House Mice Inferred Using Population Genomic Data. *Genetics*, 207(1), 297–309.
<https://doi.org/10.1534/genetics.117.300063>
- Booker, T. R., Yeaman, S., & Whitlock, M. C. (2020). Variation in recombination rate affects detection of outliers in genome scans under neutrality. *Molecular Ecology*, 29(22), 4274–4279. <https://doi.org/10.1111/mec.15501>
- Brandvain, Y., & Coop, G. (2012). Scrambling Eggs: Meiotic Drive and the Evolution of Female Recombination Rates. *Genetics*, 190(2), 709–723.
<https://doi.org/10.1534/genetics.111.136721>
- Bravo, G. A., Schmitt, C. J., & Edwards, S. V. (2021). What Have We Learned from the First 500 Avian Genomes? *Annual Review of Ecology, Evolution, and Systematics*, 52(1), 611–639. <https://doi.org/10.1146/annurev-ecolsys-012121-085928>
- Brazier, T., & Glémin, S. (2022). Diversity and determinants of recombination landscapes in flowering plants. *PLOS Genetics*, 18(8), e1010141.
<https://doi.org/10.1371/journal.pgen.1010141>
- Brekke, C., Berg, P., Gjuvsland, A. B., & Johnston, S. E. (2022). Recombination rates in pigs differ between breeds, sexes and individuals, and are associated with the RNF212, SYCP2, PRDM7, MEI1 and MSH4 loci. *Genetics Selection Evolution*, 54(1), 33.
<https://doi.org/10.1186/s12711-022-00723-9>
- Brekke, C., Johnston, S. E., Knutsen, T. M., & Berg, P. (2023). Genetic architecture of individual meiotic crossover rate and distribution in Atlantic Salmon. *Scientific Reports*, 13(1), Article 1. <https://doi.org/10.1038/s41598-023-47208-3>
- Brick, K., Thibault-Sennett, S., Smagulova, F., Lam, K.-W. G., Pu, Y., Pratto, F., Camerini-Otero, R. D., & Petukhova, G. V. (2018). Extensive sex differences at the initiation of genetic recombination. *Nature*, 561(7723), Article 7723.
<https://doi.org/10.1038/s41586-018-0492-5>
- Burri, R. (2017). Interpreting differentiation landscapes in the light of long-term linked selection. *Evolution Letters*, 1(3), 118–131. <https://doi.org/10.1002/evl3.14>
- Burt, D. W. (2002). Origin and evolution of avian microchromosomes. *Cytogenetic and*

- Genome Research*, 96(1–4), 97–112. <https://doi.org/10.1159/000063018>
- Burton, J. N., Adey, A., Patwardhan, R. P., Qiu, R., Kitzman, J. O., & Shendure, J. (2013). Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. *Nature Biotechnology*, 31(12), 1119–1125. <https://doi.org/10.1038/nbt.2727>
- Chan, A. H., Jenkins, P. A., & Song, Y. S. (2012). Genome-Wide Fine-Scale Recombination Rate Variation in *Drosophila melanogaster*. *PLOS Genetics*, 8(12), e1003090. <https://doi.org/10.1371/journal.pgen.1003090>
- Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, 4(1). <https://doi.org/10.1186/s13742-015-0047-8>
- Charlesworth, B. (1998). Measures of divergence between populations and the effect of forces that reduce variability. *Molecular Biology and Evolution*, 15(5), 538–543. <https://doi.org/10.1093/oxfordjournals.molbev.a025953>
- Charlesworth, B., & Jensen, J. D. (2021). Effects of Selection at Linked Sites on Patterns of Genetic Variability. *Annual Review of Ecology, Evolution, and Systematics*, 52(1), 177–197. <https://doi.org/10.1146/annurev-ecolsys-010621-044528>
- Charlesworth, B., Morgan, M. T., & Charlesworth, D. (1993). The effect of deleterious mutations on neutral molecular variation. *Genetics*, 134(4), 1289–1303. <https://doi.org/10.1093/genetics/134.4.1289>
- Charmantier, A., Garant, D., & Kruuk, L. E. B. (Eds.). (2014). *Quantitative genetics in the wild*. Oxford University Press.
- Coop, G., & Przeworski, M. (2007). An evolutionary view of human recombination. *Nature Reviews. Genetics*, 8(1), 23–34. <https://doi.org/10.1038/nrg1947>
- Corval, H., Cumer, T., Topaloudis, A., Roulin, A., & Goudet, J. (2023). *Where and when local adaptation happens: Lessons from the European barn owl (Tyto alba)* (p. 2023.03.17.533108). bioRxiv. <https://doi.org/10.1101/2023.03.17.533108>
- Cumer, T., Machado, A. P., Dumont, G., Bontzorlos, V., Ceccherelli, R., Charter, M.,

- Dichmann, K., Kassinis, N., Lourenço, R., Manzia, F., Martens, H.-D., Prévost, L., Rakovic, M., Roque, I., Siverio, F., Roulin, A., & Goudet, J. (2022a). Landscape and Climatic Variations Shaped Secondary Contacts amid Barn Owls of the Western Palearctic. *Molecular Biology and Evolution*, 39(1), msab343.
<https://doi.org/10.1093/molbev/msab343>
- Cumer, T., Machado, A. P., Dumont, G., Bontzorlos, V., Ceccherelli, R., Charter, M., Dichmann, K., Kassinis, N., Lourenço, R., Manzia, F., Martens, H.-D., Prévost, L., Rakovic, M., Roque, I., Siverio, F., Roulin, A., & Goudet, J. (2022b). Landscape and Climatic Variations Shaped Secondary Contacts amid Barn Owls of the Western Palearctic. *Molecular Biology and Evolution*, 39(1), msab343.
<https://doi.org/10.1093/molbev/msab343>
- Cumer, T., Machado, A. P., San-Jose, L. M., Ducrest, A.-L., Simon, C., Roulin, A., & Goudet, J. (2024). The genomic architecture of continuous plumage colour variation in the European barn owl (*Tyto alba*). *Proceedings of the Royal Society B: Biological Sciences*, 291(2014), 20231995. <https://doi.org/10.1098/rspb.2023.1995>
- Cumer, T., Machado, A. P., Siverio, F., Cherkaoui, S. I., Roque, I., Lourenço, R., Charter, M., Roulin, A., & Goudet, J. (2022). Genomic basis of insularity and ecological divergence in barn owls (*Tyto alba*) of the Canary Islands. *Heredity*, 129(5), Article 5.
<https://doi.org/10.1038/s41437-022-00562-w>
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>
- Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, 10(2), giab008.
<https://doi.org/10.1093/gigascience/giab008>
- Dimens, P. V. (2022). *pdimens/LepWrap: 4.0.1*. Zenodo.

<https://doi.org/10.5281/zenodo.6326228>

Ducrest, A.-L., Neuenschwander, S., Schmid-Siegert, E., Pagni, M., Train, C., Dylus, D., Nevers, Y., Vesztröcy, A. W., San-Jose, L. M., Dupasquier, M., Dessimoz, C., Xenarios, I., Roulin, A., & Goudet, J. (2020). New genome assembly of the barn owl (*Tyto alba alba*). *Ecology and Evolution*, *10*(5), 2284–2298.

<https://doi.org/10.1002/ece3.5991>

Eyre-Walker, A. (1993). Recombination and mammalian genome evolution. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *252*(1335), 237–243.

<https://doi.org/10.1098/rspb.1993.0071>

Fierst, J. L. (2015). Using linkage maps to correct and scaffold de novo genome assemblies: Methods, challenges, and computational tools. *Frontiers in Genetics*, *6*.

<https://www.frontiersin.org/article/10.3389/fgene.2015.00220>

Garnier, Simon, Ross, Noam, Rudis, Robert, Camargo, Pedro, A., Sciaini, Marco, Scherer, & Cédric. (2023). *viridis(Lite)—Colorblind-Friendly Color Maps for R*.

<https://doi.org/10.5281/zenodo.4679423>

Goudet, J. (2005). Hierfstat, a package for r to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, *5*(1), 184–186.

<https://doi.org/10.1111/j.1471-8286.2004.00828.x>

Grant, P. R., & Grant, B. R. (2002). Unpredictable Evolution in a 30-Year Study of Darwin's Finches. *Science*, *296*(5568), 707–711. <https://doi.org/10.1126/science.1070315>

Groenen, M. A. M., Wahlberg, P., Foglio, M., Cheng, H. H., Megens, H.-J., Crooijmans, R. P. M. A., Besnier, F., Lathrop, M., Muir, W. M., Wong, G. K.-S., Gut, I., & Andersson, L. (2009). A high-density SNP-based linkage map of the chicken genome reveals sequence features correlated with recombination rate. *Genome Research*, *19*(3), 510–519. <https://doi.org/10.1101/gr.086538.108>

Haenel, Q., Laurentino, T. G., Roesti, M., & Berner, D. (2018). Meta-analysis of chromosome-scale crossover rate variation in eukaryotes and its significance to evolutionary genomics. *Molecular Ecology*, *27*(11), 2477–2497.

<https://doi.org/10.1111/mec.14699>

Hagen, I. J., Lien, S., Billing, A. M., Elgvin, T. O., Trier, C., Niskanen, A. K., Tarka, M., Slate, J., Sætre, G.-P., & Jensen, H. (2020). A genome-wide linkage map for the house sparrow (*Passer domesticus*) provides insights into the evolutionary history of the avian genome. *Molecular Ecology Resources*, *20*(2), 544–559.

<https://doi.org/10.1111/1755-0998.13134>

Halldorsson, B. V., Palsson, G., Stefansson, O. A., Jonsson, H., Hardarson, M. T., Eggertsson, H. P., Gunnarsson, B., Oddsson, A., Halldorsson, G. H., Zink, F., Gudjonsson, S. A., Frigge, M. L., Thorleifsson, G., Sigurdsson, A., Stacey, S. N., Sulem, P., Masson, G., Helgason, A., Gudbjartsson, D. F., ... Stefansson, K. (2019). Characterizing mutagenic effects of recombination through a sequence-level genetic map. *Science*, *363*(6425), eaau1043. <https://doi.org/10.1126/science.aau1043>

Hassold, T., Hall, H., & Hunt, P. (2007). The origin of human aneuploidy: Where we have been, where we are going. *Human Molecular Genetics*, *16*(R2), R203–R208.

<https://doi.org/10.1093/hmg/ddm243>

Hewett, A. M., Stoffel, M. A., Peters, L., Johnston, S. E., & Pemberton, J. M. (2023). Selection, recombination and population history effects on runs of homozygosity (ROH) in wild red deer (*Cervus elaphus*). *Heredity*, *130*(4), Article 4.

<https://doi.org/10.1038/s41437-023-00602-z>

Hildebrand, E. M., & Dekker, J. (2020). Mechanisms and Functions of Chromosome Compartmentalization. *Trends in Biochemical Sciences*, *45*(5), 385–396.

<https://doi.org/10.1016/j.tibs.2020.01.002>

Hill, W. G., & Robertson, A. (1966). The effect of linkage on limits to artificial selection. *Genetics Research*, *8*(3), 269–294. <https://doi.org/10.1017/S0016672300010156>

Huang, Z., Xu, Z., Bai, H., Huang, Y., Kang, N., Ding, X., Liu, J., Luo, H., Yang, C., Chen, W., Guo, Q., Xue, L., Zhang, X., Xu, L., Chen, M., Fu, H., Chen, Y., Yue, Z., Fukagawa, T., ... Xu, L. (2023). Evolutionary analysis of a complete chicken genome. *Proceedings of the National Academy of Sciences*, *120*(8), e2216641120.

<https://doi.org/10.1073/pnas.2216641120>

Jerkovic', I., & Cavalli, G. (2021). Understanding 3D genome organization by multidisciplinary methods. *Nature Reviews Molecular Cell Biology*, 22(8), Article 8.

<https://doi.org/10.1038/s41580-021-00362-w>

Johnston, S. E., Béréños, C., Slate, J., & Pemberton, J. M. (2016). Conserved Genetic Architecture Underlying Individual Recombination Rate Variation in a Wild Population of Soay Sheep (*Ovis aries*). *Genetics*, 203(1), 583–598.

<https://doi.org/10.1534/genetics.115.185553>

Johnston, S. E., Huisman, J., Ellis, P. A., & Pemberton, J. M. (2017). A High-Density Linkage Map Reveals Sexual Dimorphism in Recombination Landscapes in Red Deer (*Cervus elaphus*). *G3: Genes, Genomes, Genetics*, 7(8), 2859–2870.

<https://doi.org/10.1534/g3.117.044198>

Kaur, T., & Rockman, M. V. (2014). Crossover Heterogeneity in the Absence of Hotspots in *Caenorhabditis elegans*. *Genetics*, 196(1), 137–148.

<https://doi.org/10.1534/genetics.113.158857>

Kawakami, T., Mugal, C. F., Suh, A., Nater, A., Burri, R., Smeds, L., & Ellegren, H. (2017). Whole-genome patterns of linkage disequilibrium across flycatcher populations clarify the causes and consequences of fine-scale recombination rate variation in birds.

Molecular Ecology, 26(16), 4158–4172. <https://doi.org/10.1111/mec.14197>

Kawakami, T., Smeds, L., Backström, N., Husby, A., Qvarnström, A., Mugal, C. F., Olason, P., & Ellegren, H. (2014). A high-density linkage map enables a second-generation collared flycatcher genome assembly and reveals the patterns of avian recombination rate variation and chromosomal evolution. *Molecular Ecology*, 23(16),

4035–4058. <https://doi.org/10.1111/mec.12810>

Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., & Haussler, D. (2002). The human genome browser at UCSC. *Genome Research*, 12(6), 996–1006. <https://doi.org/10.1101/gr.229102>

Kim, Y., & Nielsen, R. (2004). Linkage Disequilibrium as a Signature of Selective Sweeps.

- Genetics*, 167(3), 1513–1524. <https://doi.org/10.1534/genetics.103.025387>
- Kirkpatrick, M. (2010). How and Why Chromosome Inversions Evolve. *PLOS Biology*, 8(9), e1000501. <https://doi.org/10.1371/journal.pbio.1000501>
- Koehler, K. E., Hawley, R. S., Sherman, S., & Hassold, T. (1996). Recombination and nondisjunction in humans and flies. *Hum Mol Genet*, 5 Spec No, 1495–1504. https://doi.org/10.1093/hmg/5.supplement_1.1495
- Kong, A., Barnard, J., Gudbjartsson, D. F., Thorleifsson, G., Jonsdottir, G., Sigurdardottir, S., Richardsson, B., Jonsdottir, J., Thorgeirsson, T., Frigge, M. L., Lamb, N. E., Sherman, S., Gulcher, J. R., & Stefansson, K. (2004). Recombination rate and reproductive success in humans. *Nature Genetics*, 36(11), 1203–1206. <https://doi.org/10.1038/ng1445>
- Kong, A., Gudbjartsson, D. F., Sainz, J., Jonsdottir, G. M., Gudjonsson, S. A., Richardsson, B., Sigurdardottir, S., Barnard, J., Hallbeck, B., Masson, G., Shlien, A., Palsson, S. T., Frigge, M. L., Thorgeirsson, T. E., Gulcher, J. R., & Stefansson, K. (2002). A high-resolution recombination map of the human genome. *Nature Genetics*, 31(3), Article 3. <https://doi.org/10.1038/ng917>
- Kong, A., Masson, G., Frigge, M. L., Gylfason, A., Zusmanovich, P., Thorleifsson, G., Olason, P. I., Ingason, A., Steinberg, S., Rafnar, T., Sulem, P., Mouy, M., Jonsson, F., Thorsteinsdottir, U., Gudbjartsson, D. F., Stefansson, H., & Stefansson, K. (2008). Detection of sharing by descent, long-range phasing and haplotype imputation. *Nature Genetics*, 40(9), 1068–1075. <https://doi.org/10.1038/ng.216>
- Kong, A., Thorleifsson, G., Gudbjartsson, D. F., Masson, G., Sigurdsson, A., Jonasdottir, A., Walters, G. B., Jonasdottir, A., Gylfason, A., Kristinsson, K. T., Gudjonsson, S. A., Frigge, M. L., Helgason, A., Thorsteinsdottir, U., & Stefansson, K. (2010). Fine-scale recombination rate differences between sexes, populations and individuals. *Nature*, 467(7319), Article 7319. <https://doi.org/10.1038/nature09525>
- Lack, D., & Lack, E. (1958). The Nesting of the Long-Tailed Tit. *Bird Study*, 5(1), 1–19. <https://doi.org/10.1080/00063655809475897>

- Lam, I., & Keeney, S. (2014). Mechanism and Regulation of Meiotic Recombination Initiation. *Cold Spring Harbor Perspectives in Biology*, 7(1), a016634.
<https://doi.org/10.1101/cshperspect.a016634>
- Lam, I., & Keeney, S. (2015). Nonparadoxical evolutionary stability of the recombination initiation landscape in yeast. *Science*, 350(6263), 932–937.
<https://doi.org/10.1126/science.aad0814>
- Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv:1303.3997 [q-Bio]*. <http://arxiv.org/abs/1303.3997>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079.
<https://doi.org/10.1093/bioinformatics/btp352>
- Li, N., & Stephens, M. (2003). Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. *Genetics*, 165(4), 2213–2233.
- Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B. R., Sabo, P. J., Dorschner, M. O., Sandstrom, R., Bernstein, B., Bender, M. A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L. A., Lander, E. S., & Dekker, J. (2009). Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science*, 326(5950), 289–293. <https://doi.org/10.1126/science.1181369>
- Machado, A. P., Cumer, T., Iseli, C., Beaudoin, E., Ducrest, A.-L., Dupasquier, M., Guex, N., Dichmann, K., Lourenço, R., Lusby, J., Martens, H.-D., Prévost, L., Ramsden, D., Roulin, A., & Goudet, J. (2022). Unexpected post-glacial colonisation route explains the white colour of barn owls (*Tyto alba*) from the British Isles. *Molecular Ecology*, 31(2), 482–497. <https://doi.org/10.1111/mec.16250>
- Machado, A. P., Topaloudis, A., Cumer, T., Lavanchy, E., Bontzorlos, V., Ceccherelli, R., Charter, M., Kassinis, N., Lymberakis, P., Manzia, F., Ducrest, A.-L., Dupasquier, M., Guex, N., Roulin, A., & Goudet, J. (2022). Genomic consequences of colonisation,

- migration and genetic drift in barn owl insular populations of the eastern Mediterranean. *Molecular Ecology*, 31(5), 1375–1388.
<https://doi.org/10.1111/mec.16324>
- Marx, V. (2023). Method of the year: Long-read sequencing. *Nature Methods*, 20(1), Article 1. <https://doi.org/10.1038/s41592-022-01730-w>
- McAuley, J. B., Servin, B., Peters, L., Hagen, I. J., Niskanen, A. K., Ringsby, T. H., Husby, A., Jensen, H., & Johnston, S. E. (2023). *Sex-differences in the genetic architecture of individual recombination rates in wild house sparrows (Passer domesticus)*. bioRxiv.
<https://doi.org/10.1101/2023.01.26.525019>
- McVean, G. A. T., Myers, S. R., Hunt, S., Deloukas, P., Bentley, D. R., & Donnelly, P. (2004). The Fine-Scale Structure of Recombination Rate Variation in the Human Genome. *Science*, 304(5670), 581–584. <https://doi.org/10.1126/science.1092500>
- Myers, S., Bottolo, L., Freeman, C., McVean, G., & Donnelly, P. (2005). A Fine-Scale Map of Recombination Rates and Hotspots Across the Human Genome. *Science*, 310(5746), 321–324. <https://doi.org/10.1126/science.1117196>
- Myers, S., Bowden, R., Tumian, A., Bontrop, R. E., Freeman, C., MacFie, T. S., McVean, G., & Donnelly, P. (2010). Drive Against Hotspot Motifs in Primates Implicates the PRDM9 Gene in Meiotic Recombination. *Science*, 327(5967), 876–879.
<https://doi.org/10.1126/science.1182363>
- O'Reilly, P. F., Birney, E., & Balding, D. J. (2008). Confounding between recombination and selection, and the Ped/Pop method for detecting selection. *Genome Research*, 18(8), 1304–1313. <https://doi.org/10.1101/gr.067181.107>
- Otto, S. P., & Lenormand, T. (2002). Resolving the paradox of sex and recombination. *Nature Reviews Genetics*, 3(4), Article 4. <https://doi.org/10.1038/nrg761>
- Otto, S. P., & Payseur, B. A. (2019). Crossover Interference: Shedding Light on the Evolution of Recombination. *Annual Review of Genetics*, 53(1), 19–44.
<https://doi.org/10.1146/annurev-genet-040119-093957>
- Pemberton, J. m. (2008). Wild pedigrees: The way forward. *Proceedings of the Royal*

Society B: Biological Sciences, 275(1635), 613–621.

<https://doi.org/10.1098/rspb.2007.1531>

Pemberton, T. J., Absher, D., Feldman, M. W., Myers, R. M., Rosenberg, N. A., & Li, J. Z. (2012). Genomic Patterns of Homozygosity in Worldwide Human Populations. *The American Journal of Human Genetics*, 91(2), 275–292.

<https://doi.org/10.1016/j.ajhg.2012.06.014>

Peñalba, J. V., Deng, Y., Fang, Q., Joseph, L., Moritz, C., & Cockburn, A. (2020). Genome of an iconic Australian bird: High-quality assembly and linkage map of the superb fairy-wren (*Malurus cyaneus*). *Molecular Ecology Resources*, 20(2), 560–578.

<https://doi.org/10.1111/1755-0998.13124>

Peñalba, J. V., & Wolf, J. B. W. (2020). From molecules to populations: Appreciating and estimating recombination rate variation. *Nature Reviews Genetics*, 21(8), 476–492.

<https://doi.org/10.1038/s41576-020-0240-1>

Peona, V., Weissensteiner, M. H., & Suh, A. (2018). How complete are “complete” genome assemblies?—An avian perspective. *Molecular Ecology Resources*, 18(6),

1188–1195. <https://doi.org/10.1111/1755-0998.12933>

Phillips, D., Jenkins, G., Macaulay, M., Nibau, C., Wnetrzak, J., Fallding, D., Colas, I., Oakey, H., Waugh, R., & Ramsay, L. (2015). The effect of temperature on the male and female recombination landscape of barley. *New Phytologist*, 208(2), 421–429.

<https://doi.org/10.1111/nph.13548>

Posit team. (2022). *RStudio: Integrated Development Environment for R*. Posit Software, PBC. <http://www.posit.co/>

Pya, N., & Wood, S. N. (2015). Shape constrained additive models. *Statistics and Computing*, 25(3), 543–559. <https://doi.org/10.1007/s11222-013-9448-7>

Quinlan, A. R. (2014). BEDTools: The Swiss-Army Tool for Genome Feature Analysis. *Current Protocols in Bioinformatics*, 47(1), 11.12.1-11.12.34.

<https://doi.org/10.1002/0471250953.bi11112s47>

R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. R

- Foundation for Statistical Computing. <https://www.R-project.org/>
- Rastas, P. (2017). Lep-MAP3: Robust linkage mapping even for low-coverage whole genome sequencing data. *Bioinformatics*, *33*(23), 3726–3732.
<https://doi.org/10.1093/bioinformatics/btx494>
- Roulin, A., Müller, W., Sasvári, L., Dijkstra, C., Ducrest, A.-L., Riols, C., Wink, M., & Lubjuhn, T. (2004). Extra-pair paternity, testes size and testosterone level in relation to colour polymorphism in the barn owl *Tyto alba*. *Journal of Avian Biology*, *35*(6), 492–500.
<https://doi.org/10.1111/j.0908-8857.2004.03294.x>
- Sardell, J. M., & Kirkpatrick, M. (2019). Sex Differences in the Recombination Landscape. *The American Naturalist*, *195*(2), 361–379. <https://doi.org/10.1086/704943>
- Schild, D. R., Pasquesi, G. I. M., Perry, B. W., Adams, R. H., Nikolakis, Z. L., Westfall, A. K., Orton, R. W., Meik, J. M., Mackessy, S. P., & Castoe, T. A. (2020). Snake Recombination Landscapes Are Concentrated in Functional Regions despite PRDM9. *Molecular Biology and Evolution*, *37*(5), 1272–1294.
<https://doi.org/10.1093/molbev/msaa003>
- Shakoori, A. R. (2017). Fluorescence In Situ Hybridization (FISH) and Its Applications. *Chromosome Structure and Aberrations*, 343–367.
https://doi.org/10.1007/978-81-322-3673-3_16
- Shanfelter, A. F., Archambeault, S. L., & White, M. A. (2019). Divergent Fine-Scale Recombination Landscapes between a Freshwater and Marine Population of Threespine Stickleback Fish. *Genome Biology and Evolution*, *11*(6), 1552–1572.
<https://doi.org/10.1093/gbe/evz090>
- Sheldon, B. C., Kruuk, L. E. B., & Alberts, S. C. (2022). The expanding value of long-term studies of individuals in the wild. *Nature Ecology & Evolution*, *6*(12), Article 12.
<https://doi.org/10.1038/s41559-022-01940-7>
- Signorell, A. (2023). *DescTools: Tools for Descriptive Statistics*.
<https://andrisignorell.github.io/DescTools/>
- Singhal, S., Leffler, E. M., Sannareddy, K., Turner, I., Venn, O., Hooper, D. M., Strand, A. I.,

- Li, Q., Raney, B., Balakrishnan, C. N., Griffith, S. C., McVean, G., & Przeworski, M. (2015). Stable recombination hotspots in birds. *Science*, *350*(6263), 928–932. <https://doi.org/10.1126/science.aad0843>
- Sinnwell, J. P., Therneau, T. M., & Schaid, D. J. (2014). The kinship2 R Package for Pedigree Data. *Human Heredity*, *78*(2), 91–93. <https://doi.org/10.1159/000363105>
- Smeds, L., Qvarnström, A., & Ellegren, H. (2016). Direct estimate of the rate of germline mutation in a bird. *Genome Research*, *26*(9), 1211–1218. <https://doi.org/10.1101/gr.204669.116>
- Smukowski Heil, C. S., Ellison, C., Dubin, M., & Noor, M. A. F. (2015). Recombining without Hotspots: A Comprehensive Evolutionary Portrait of Recombination in Two Closely Related Species of *Drosophila*. *Genome Biology and Evolution*, *7*(10), 2829–2842. <https://doi.org/10.1093/gbe/evv182>
- Spence, J. P., & Song, Y. S. (2019). Inference and analysis of population-specific fine-scale recombination maps across 26 diverse human populations. *Science Advances*, *5*(10), eaaw9206. <https://doi.org/10.1126/sciadv.aaw9206>
- Stapley, J., Feulner, P. G. D., Johnston, S. E., Santure, A. W., & Smadja, C. M. (2017). Variation in recombination frequency and distribution across eukaryotes: Patterns and processes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1736), 20160455. <https://doi.org/10.1098/rstb.2016.0455>
- Tease, C., & Hultén, M. A. (2004). Inter-sex variation in synaptonemal complex lengths largely determine the different recombination rates in male and female germ cells. *Cytogenetic and Genome Research*, *107*(3–4), 208–215. <https://doi.org/10.1159/000080599>
- Tennekes, M. (2018). tmap: Thematic Maps in R. *Journal of Statistical Software*, *84*(6), 1–39. <https://doi.org/10.18637/jss.v084.i06>
- Terhorst, J., Kamm, J. A., & Song, Y. S. (2017). Robust and scalable inference of population history from hundreds of unphased whole genomes. *Nature Genetics*, *49*(2), Article 2. <https://doi.org/10.1038/ng.3748>

- van Oers, K., Santure, A. W., De Cauwer, I., van Bers, N. E., Crooijmans, R. P., Sheldon, B. C., Visser, M. E., Slate, J., & Groenen, M. A. (2014). Replicated high-density genetic maps of two great tit populations reveal fine-scale genomic departures from sex-equal recombination rates. *Heredity*, *112*(3), Article 3.
<https://doi.org/10.1038/hdy.2013.107>
- Veller, C., Kleckner, N., & Nowak, M. A. (2019). A rigorous measure of genome-wide genetic shuffling that takes into account crossover positions and Mendel's second law. *Proceedings of the National Academy of Sciences*, *116*(5), 1659–1668.
<https://doi.org/10.1073/pnas.1817482116>
- Wall, J. D., Robinson, J. A., & Cox, L. A. (2022). High-Resolution Estimates of Crossover and Noncrossover Recombination from a Captive Baboon Colony. *Genome Biology and Evolution*, *14*(4), evac040. <https://doi.org/10.1093/gbe/evac040>
- Waters, P. D., Patel, H. R., Ruiz-Herrera, A., Álvarez-González, L., Lister, N. C., Simakov, O., Ezaz, T., Kaur, P., Frere, C., Grützner, F., Georges, A., & Graves, J. A. M. (2021). Microchromosomes are building blocks of bird, reptile, and mammal chromosomes. *Proceedings of the National Academy of Sciences*, *118*(45), e2112494118.
<https://doi.org/10.1073/pnas.2112494118>
- Wei, T., & Simko, V. (2021). *R package "corrplot": Visualization of a Correlation Matrix*.
<https://github.com/taiyun/corrplot>
- Weir, B. S., & Goudet, J. (2017). A Unified Characterization of Population Structure and Relatedness. *Genetics*, *206*(4), 2085–2103.
<https://doi.org/10.1534/genetics.116.198424>
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., ... Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686.
<https://doi.org/10.21105/joss.01686>
- Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood

estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(1), 3–36.

<https://doi.org/10.1111/j.1467-9868.2010.00749.x>

Zelkowski, M., Olson, M. A., Wang, M., & Pawlowski, W. (2019). Diversity and Determinants of Meiotic Recombination Landscapes. *Trends in Genetics*, 35(5), 359–370.

<https://doi.org/10.1016/j.tig.2019.02.002>

Zhang, H., Lundberg, M., Ponnikas, S., Hasselquist, D., & Hansson, B. (2023). *Male-biased recombination at chromosome ends in a songbird revealed by precisely mapping crossover positions* (p. 2023.12.19.572321). bioRxiv.

<https://doi.org/10.1101/2023.12.19.572321>

Zickler, D., & Kleckner, N. (2015). Recombination, Pairing, and Synapsis of Homologs during Meiosis. *Cold Spring Harbor Perspectives in Biology*, 7(6), a016626.

<https://doi.org/10.1101/cshperspect.a016626>

ANNEX II

Authors: Tristan Cumer^{1,2}, Eléonore Lavanchy^{1,2}, Jérôme Goudet^{1,2}

¹ Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland

² Swiss Institute of Bioinformatics, University of Lausanne, Lausanne, Switzerland

This annex provides information about the phasing of the 502 owls described as the “Reference panel” used in chapters III and IV. The first part contains a methodological description of the phasing while the second part contains an evaluation of the phasing.

Authors contributions: EL performed the read-based phasing and wrote the methodological description section. TC performed the statistical phasing, evaluated both phasing and wrote the section on the evaluation of phasing. All co-authors gave feedbacks on the writing.

Phasing of the Reference Panel

Phasing of the SNPs

The sampling, sequencing and genotyping of the SNPs are described in Annex I. The set of filtered variants was phased in two steps. First, individual variants were phased using a read-based approach in which reads covering multiple heterozygous sites were used to resolve local haplotypes. To do so, `WhatsHap` v1.4 [3] was run independently for each individual. We considered the mean recombination rate (estimated from maps described in Annex I) when phasing (`--recombrate 2`). All individuals belonging to one (or more) trios were phased with `WhatsHap` pedigree-mode. Parents which were present in more than one trios (which have more than one kid) were phased along with the kid who had the higher coverage.

After the read-based phasing, we performed an additional step of filtering. We kept only bi-allelic SNPs, sites with minor allele count (MAC) > 5 and missing data $< 5\%$. MAC and missingness were estimated on the set of unrelated individuals (with pairwise allele-sharing kinship < 0.03125 which resulted in 187 individuals). After filtering, we ended up with 10,451,268 variants.

Filtered variants were then statistically phased with `Shape-It` v4.1.2 [2]. This algorithm integrates local individual phase and applies an approach based on coalescence and recombination to statistically phase haplotypes and impute missing data. Using `Shape-It`, we phased the 187 unrelated and 315 related individuals separately. The 187 unrelated individuals were phased together. Then each related individual was phased with the unrelated set separately to avoid any effect of family structure in the phasing panel. `Shape-It` was run following the manual's instructions for better accuracy: the number of conditioning neighbors in the PBWT was set to 8, and the MCMC chain was run with 10 burn-in generations, 5 pruning iterations, each separated by 1 burn-in iteration, and 10 main iterations.

Evaluation of the phasing

To assess the quality of the phasing, we examined phase accuracy by using the switch-error-rate metric [1]. When comparing two phasings for an individual's variants, a switch error occurs when a heterozygous site has its phase switched relative to that of the previous

heterozygous site. Thus, for each individual, we compared the true local phasing, inferred either from the read-based approach or the trio approach (both from `WhatsHap`) and the statistical phasing of this individual's variants statistically phased by `Shape-IT`, with read-based/trio phase information ignored for the individual considered only (same version and parameters described in the previous paragraph). The final estimation of the switch error rate was done using the `switchError` code to compare both phasing sets (available on [O. Delaneau GitHub](#)) and is presented in figure 1.

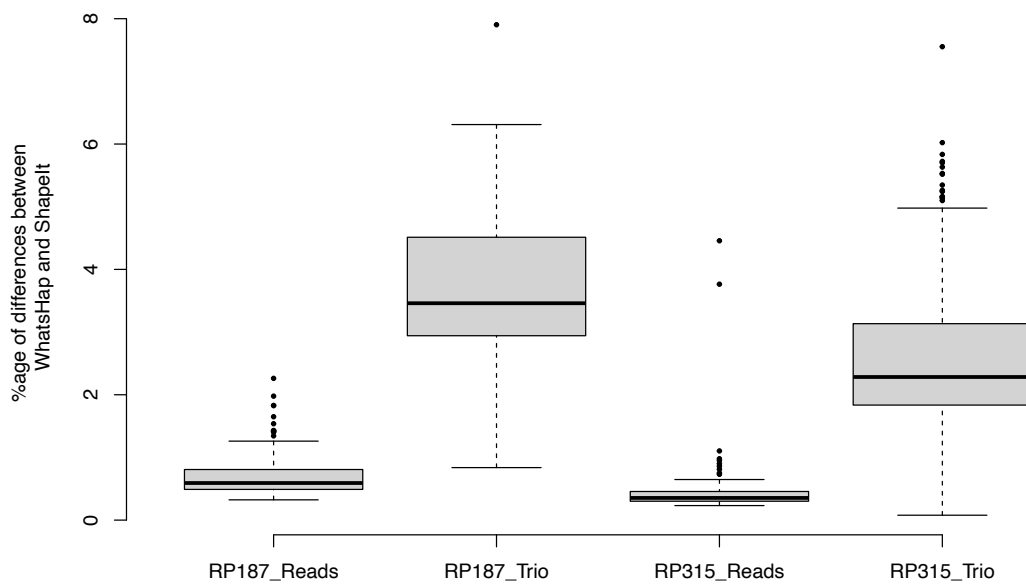


Figure 1: Percentage of difference between `WhatsHap` and `Shape-IT` phasing.

The difference is bigger for individuals phased with trios but we show in figure 2 that it does not mean that trio-phased individuals are phased wrongly. We show in figure 2 that individuals phased with trios have longer phased-blocks and more SNPs in each block.

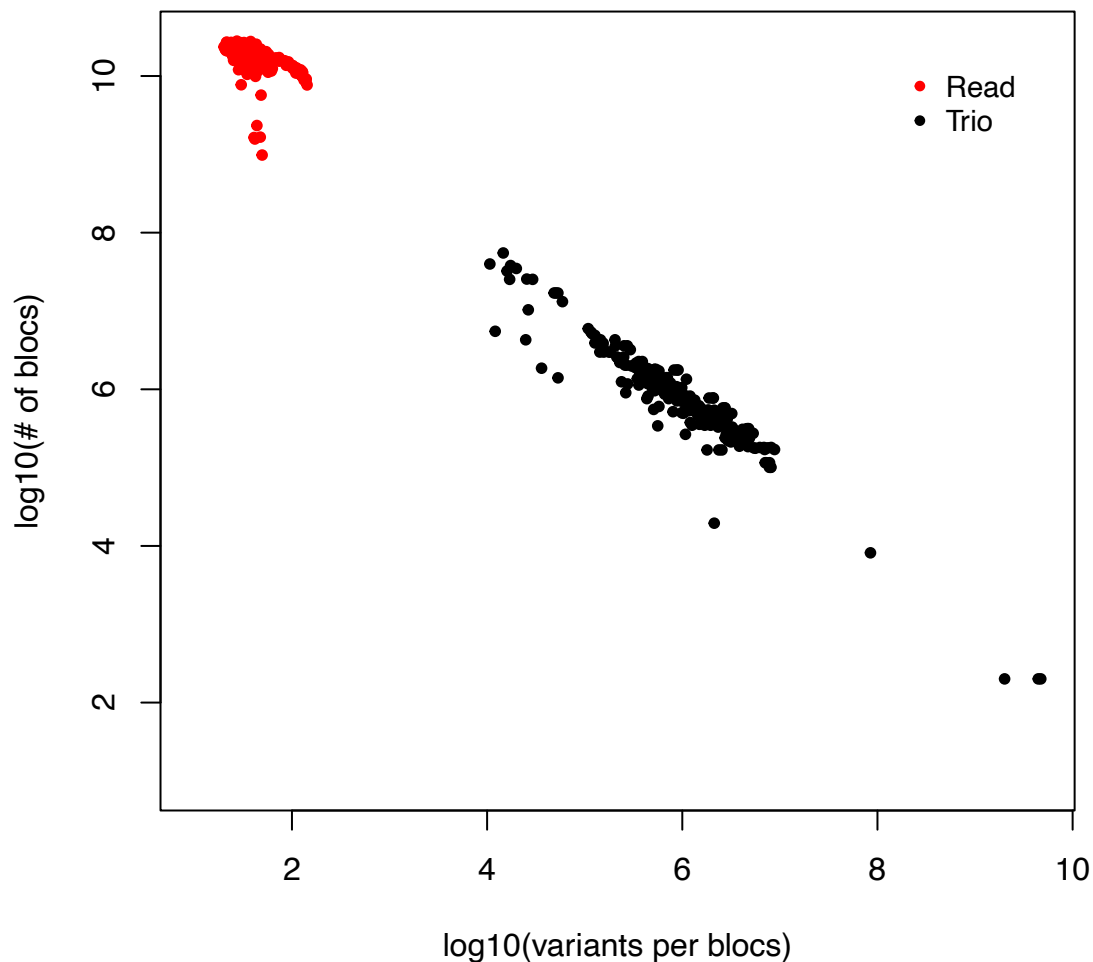


Figure 2: number of phased blocks and number of variants per blocks in individuals phased with read-based only and individuals phased with trios.

References

- [1] Y. Choi, A.P. Chan, E. Kirkness, A. Telenti, and N.J. Schork. Comparison of phasing strategies for whole human genomes. *PLOS Genetics*, 14(4):e1007308, 2018.
- [2] O. Delaneau, J-F. Zagury, M.R. Robinson, J.L. Marchini, and E.T. Dermitzakis. Accurate, scalable and integrative haplotype estimation. *Nature Communications*, 10(1):5436, 2019.
- [3] M. Martin, M. Patterson, S. Garg, S.O. Fischer, N. Pisanti, G.W. Klau, A. Schöenhuth, and T. Marschall. *Whatshap*: fast and accurate read-based phasing. *bioRxiv*, 2016.

ANNEX III

Authors : Tristan Cumer^{1,2}, Eléonore Lavanchy^{1,2}, Alexandros Topaloudis^{1,2}, Anne-Lyse Ducrest^{1,2}, Céline Simon^{1,2}, Alexandre Roulin^{1,2}, Jérôme Goudet^{1,2}

¹ Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland

² Swiss Institute of Bioinformatics, University of Lausanne, Lausanne, Switzerland

This annex describes the low-coverage data generation for the 2,765 owls used in chapter III. It includes the sampling, the sequencing, the genotyping as well as the validation of our low-coverage procedure with 32 individuals we sequenced at both high and low coverage.

Authors contributions: the sampling, genotyping and validation were performed by TC, EL and AT. The DNA extraction and sample preparation were performed by A-LD and CS. EL, TC, AT and A-LD wrote the text and all co-authors provided feedbacks on each section.

Sampling

We sequenced 2,820 owls from the pedigree population (fractions of the pedigree population sampled for sequencing per year are presented in figure 1). Samples were collected between 1986 and 2020 and we prioritized individuals with complete families that reproduced multiple times and for which phenotypic data were available. We also maximised the sampling of inbred individuals (based on the pedigree). Blood samples of the individuals were kept in a -80 degrees Celsius freezer since the time of sampling.

For verification of sequencing performance we included 32 individuals that had already been sequenced at high coverage in the Reference Panel (described in Annex I & II). Ten of these individuals were sequenced three times each in low coverage, once per flow cell to quantify potential batch effects. We refer to these individuals as ‘triplicates’. The 22 remaining samples used to verify the quality of low-coverage genotyping are called ‘duplicates’.

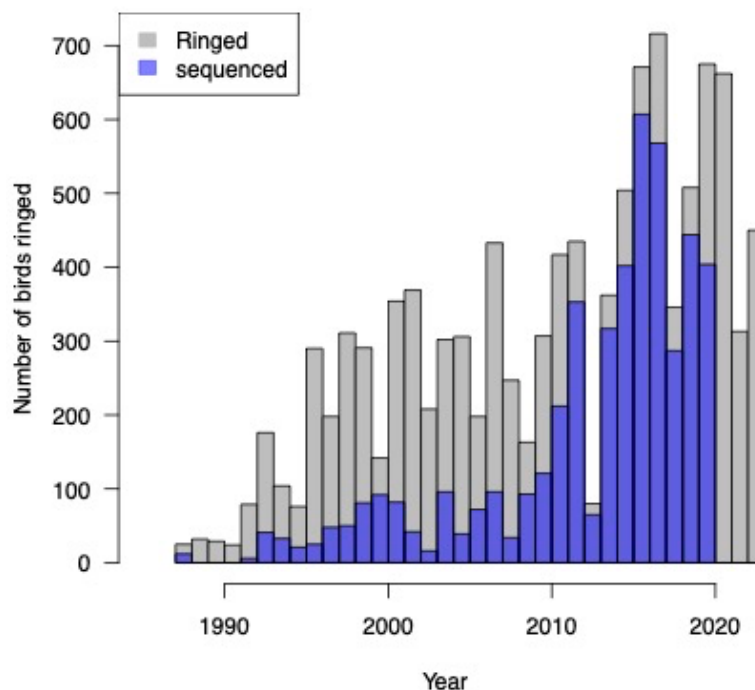


Figure 1: number of individuals (in gray) and number of individuals we sequenced (in blue) in the Swiss barn owl population per year. Year corresponds to ringing year (which differs from birth year for birds which were not born in our study area). Sampling stops in 2020 because the project started in 2021 when the field season was not over yet.

Blood was collected from each barn owl's brachial vein and stored in liquid nitrogen before being transferred to -80°C freezers. Genomic DNA was extracted from blood samples using the DNeasy Tissue Kit (Qiagen, Switzerland) and the Biosprint robot 96 (Qiagen, Switzerland) and stored at -20°C . We randomized the position of the DNA samples of the 2,820 owls on 30 different 96-well plates. Every plate included two empty wells as a measure of contamination control.

The DNA was quantified using Quant-it PicoGreen dsDNA Assay kit (Thermo Scientific, Switzerland). The DNA was diluted in 10 mM Tris-HCl to 1.5 to 2.5ng/ μl . The libraries were prepared with plexWell 96 kits (SeqWell, USA) and sequenced using NovaSeq 6000 (Illumina) at the genomic technologies facility of the University of Lausanne (GTF). After sequencing the samples on 3 different flow cells, the average coverage was 1.91X.

Raw reads processing

Raw reads were trimmed with `Trimmomatic v.0.36` [1] and aligned to the reference barn owl genome [6] using `BWA-MEM v.0.7.15` [5]. `Qualibam v2.2.1` was then used to extract the mean coverage and the coverage of the Super-Scaffold 13 (belonging to the Z chromosome [6]).

Sexing of the low coverage individuals

Since males are diploid and females are haploid for the sexual (Z) chromosome, the ratio of mean coverages at the sexual Z chromosome (Super-Scaffold 13) over the autosomal region allowed us to infer the sex of individuals. We identified two distinct groups of individuals, with a ratio close to one for males and 0.5 for females (figure 2).

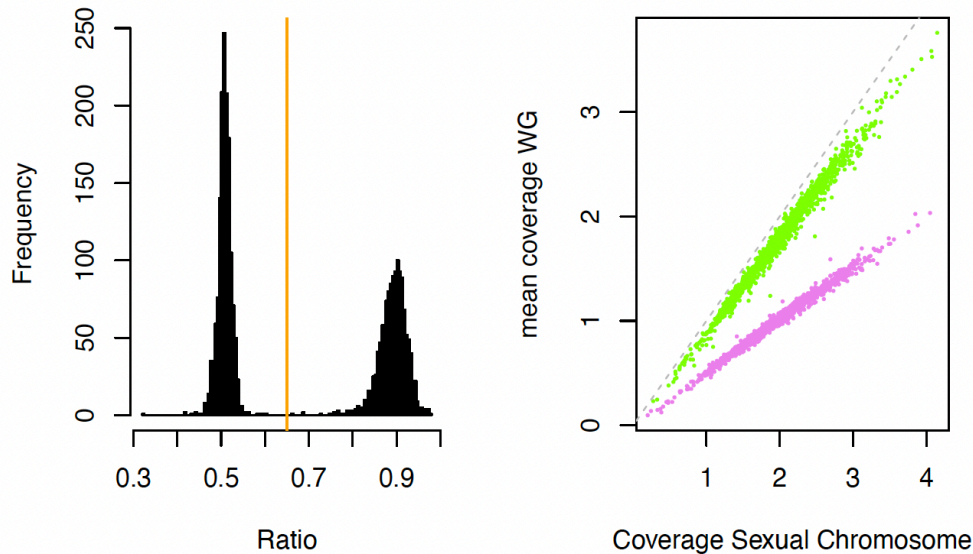


Figure 2: Sexing of the individuals based on whole genome sequencing. Left panel: Histogram of the ratio between the coverage of the sexual chromosome and the mean coverage along the genome. The vertical line at 0.65 depicts the threshold used to classify individuals as males or females. Right panel: plot of the mean coverage of the whole genome (WG) as a function of the mean coverage along the sexual chromosome. Individuals with a ratio lower than 0.65 were considered females (ZW, in pink), while individuals with a coverage higher are considered as males (WW, in green).

Genotyping

Description of the method

We employed the *GLIMPSE* [7] software for the imputation and phasing of our low coverage individuals. *GLIMPSE* phases and imputes low coverage genotypes based on high coverage haplotypes (called the reference panel) and can only be applied on bi-allelic SNPs. We used the 502 owls sequenced at high coverage and described in Annexes I and II as the reference panel. We genotyped low-coverage individuals at the variant sites (and for the alleles) present in the reference panel only. We used the first version of *GLIMPSE* (v1.1.1) because it is more suitable for small reference panels (as indicated on the [GLIMPSE website](#)). Briefly, the *GLIMPSE* pipeline consists of four steps.

The first step is called chunking (implemented in the *GLIMPSE_chunk* method) and consists of splitting the chromosome into chunks for efficient imputation and phasing. As chunks become smaller, the speed of the method increases, but the accuracy decreases. Based on both the amount of missing data and chromosome length, the *GLIMPSE_chunk* method is used to automatically estimate the best chunk size and split the chromosome.

The second step of the pipeline consists of the phasing and imputation (implemented in the *GLIMPSE_phase* method) of the genomic chunks. Iterative improvements are made to low-coverage genotypes' likelihoods and phasing for each individual separately. In each iteration, *GLIMPSE* uses a slightly modified version of *Shape-It* v.4 for statistical phasing of the target low-coverage sample (using both the haplotypes present in the reference panel as well as other low coverage haplotypes). Next, the closest haplotypes (defined as the haplotypes sharing the highest amount of identical-by-descent (IBD) segments with the target haplotype) are identified from the reference panel and the other low-coverage imputed haplotypes with the positional Burrows–Wheeler transform (PBWT) algorithm. These closest haplotypes are then used for imputation of the target haplotype. The imputation is performed using a Hidden Markov Model (HMM) in which individual genotypes consist of two hidden states: the reference or the alternate allele. Emission probabilities depend on the alleles of the selected closest haplotypes, the target sample's posterior genotype likelihood (from previous iterations) and an error rate. The transition probabilities depend on the alleles of the selected closest haplotypes and the recombination probability. Within diploid individuals, each haplotype is imputed separately. The new iteration begins once the imputation has been completed.

Once the phasing and imputation of the chunks are completed, the next step is to merge the different chunks without losing the phase information. This step is called ligation and is implemented in the *GLIMPSE_ligate* method. Ligation of adjacent chunks without losing phase information is possible because the chunks are slightly overlapping.

The last step of the *GLIMPSE* pipeline is the haplotype identification (implemented in the *GLIMPE_sample* method). It identifies the most likely haplotype based on posterior genotypes' likelihood and phase information.

Description of the parameters we used

We followed the [GLIMPSE tutorial](#) for genotyping, imputation and phasing of the 2,768 unique individuals sequenced at low coverage. The tutorial provides a detailed explanation and examples of all the commands needed for genotyping, phasing, and imputation. The code we used for the owls is available on [GitHub](#).

To extract the variant positions as well as their corresponding alleles in the reference panel, we used the query method from *BCFTOOLS* v.1.15.1 [2]. We then used *BCFTOOLS*

mpileup and *call* methods [4] to determine low coverage individuals' genotypes likelihood at each variant position identified in the previous step (with the *-T* and *-C allele* parameters).

We performed chunking with the *GLIMPSE_chunk* method with default parameters (*--window-size 2000000* and *--buffer-size 200000*). Next, we conducted phasing and imputation with the *GLIMPSE_phase* method, with the maximum number of iterations (*--burnin 100* and *--main 15*) and with a large effective population size parameter value (*--ne 10000*). If a map was available, we took into account the recombination landscape (with the *--map* parameter) by including the recombination maps from Annex I. All individuals were considered diploid for the genotyping of the autosomes (see Annex IV for a precise description of the sequencing and genotyping of the sexual chromosome). We then ligated the different chunks with the *GLIMPSE_ligate* method and extracted the most likely haplotype with the *GLIMPSE_sample* method and the *--solve* parameter.

As mentioned in the sampling section of this Annex, we selected 32 individuals which we sequenced at both high and low coverage in order to ensure that the genotypes we obtained with the *GLIMPSE* pipeline were reliable. Among these samples, 10 were sequenced three times at low coverage: once per flow-cell to control for potential batch effect and are herein referred to as 'triplicates'. The 22 remaining individuals were sequenced once at low coverage and once at high coverage and are referred to as 'duplicates'. As mentioned above, *GLIMPSE* uses the haplotypes with the higher fraction of IBD segments with the target sample for both phasing and imputation. Consequently, having the same individual in the reference panel and as a target for imputation would lead to biased results. Thus, we ran the *GLIMPSE* pipeline five times: once with the full reference panel (the 502 individuals) and with the 2,768 unique low-coverage samples. This run generated the 'main' data of the 3K owls: namely the data used in all subsequent analyses. We then removed the 22 duplicates from the reference panel, included them in the low-coverage set and ran the *GLIMPSE* pipeline again (with a reference panel of 480 individuals and a low coverage data set of 2,790 individuals). Since *GLIMPSE* also samples the closest haplotypes from the rest of the low-coverage set, we also ran *GLIMPSE* separately for the three sequencing of the triplicates (with a reference panel of 492 individuals and a low coverage data set of 2,778 individuals, called runs a, b and c).

Validation of the genotyping

After genotyping, we randomly selected a subset of 50,000 SNPs to assess the correlation between dosages among the triplicates runs of GLIMPSE (runs a, b, and c). The correlations are shown in figure 3. Individual correlations are all above 0.99, indicating that: i) there is no batch effect in our sequencing design, and ii) there is high concordance between the GLIMPSE runs.

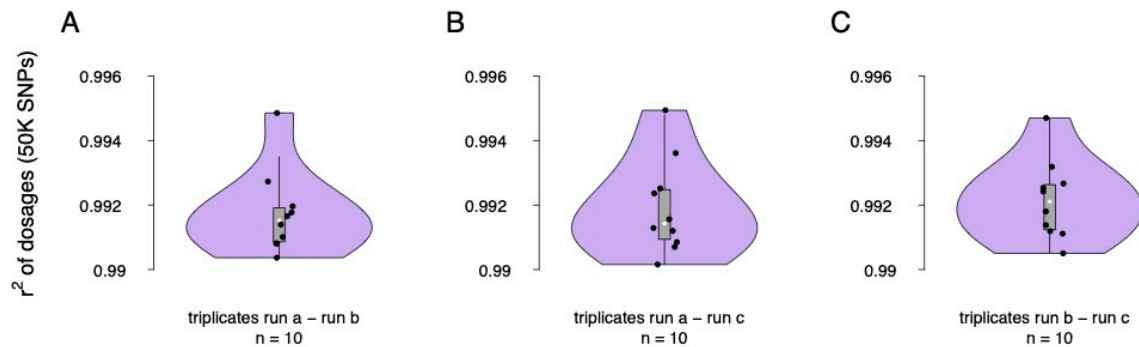


Figure 3: correlation of dosages (from 50,000 SNPs) between the different runs of GLIMPSE including the triplicates in the low-coverage set. Panel A contrasts run a and run b. Panel B contrasts run a and run c and panel C contrasts run b and run c. Each point represents the correlation between two runs for one individual.

We used the same subset of 50,000 SNPs to assess the correlation of dosages between high and low coverages. This assessment was conducted on 32 individuals (duplicates and triplicates together) and is shown in figure 4. Since we found virtually no difference between the dosages from the three runs of the triplicates, we selected genotypes from run a. There is a strong correlation, all above 0.94, between genotype dosage from high and low coverage, indicating that our low-coverage genotypes are reliable.

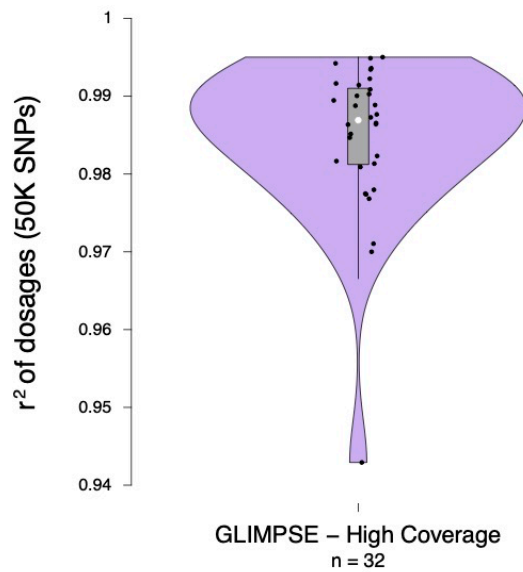


Figure 4: correlation of dosages (from 50,000 SNPs) for individuals sequenced at both high and low coverages. Each point represents the correlation for one individual.

Next, we compare the kinship between each pair of individuals estimated from the pedigree and from genomic data (figure 5). The pedigree-based kinship was estimated using the function *kinship* from the *kinship2* R package [8]. The genomic-based kinship was estimated with the *beta.dosage()* function from the *hierfstat* R package [3]. Overall, there is high consistency between both approaches.

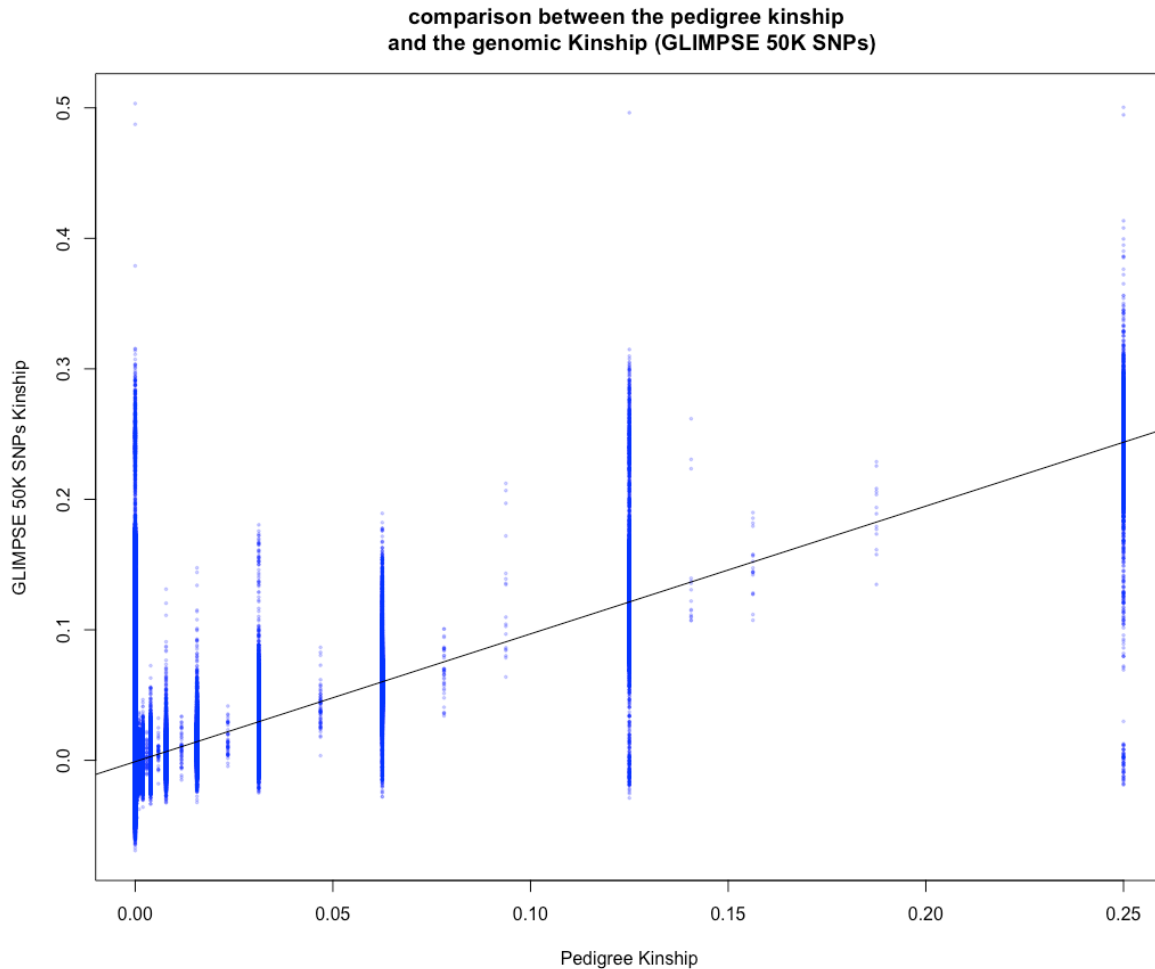


Figure 5: Comparison of the genomic-based kinship estimated from GLIMPSE and 50,000 SNPS and the pedigree based kinship. Each point represents a pair of individuals.

Due to Mendelian segregation and recombination, it is expected that genomic-based kinship will randomly vary around pedigree-based kinship [9]. Nevertheless, large differences between both approaches are not expected and could be caused by two different factors. One of the reasons is that the pedigree is incorrect or incomplete. Due to the fact that this population is a natural population, we cannot be certain that the parents we assigned to each offspring are their true parents. For the female, this is less likely than for the male, since the female is incubating the eggs for quite some time. Generally, when a male is seen visiting a nestbox, we assume that he is the father, but it is possible that he may simply be another prospecting male. Additionally, it is possible that some offspring are the result of extra-pair copulations. Another explanation for large differences between both kinships is that we did not sequence the individual we thought we sequenced. For purely genomic analyses, this is not a problem. However, it is a problem when phenotypic data are linked to genomic data. Therefore, we manually examined each pair of individuals

for which the difference between genomic and pedigree kinship was too high (documented in the following section).

Samples removal

We removed 5 pairs of samples because they were genomic twins. We also removed two additional individuals because they had weirdly high kinship with all the other individuals in the data set.

Comparing genomic and pedigree kinship

As mentioned above, we don't expect big discrepancies between genomic and pedigree-based kinship. We plot the subtraction of the genomic-based kinship to the pedigree kinship for each pair of individuals in figure 6 (mean difference = 0.001).

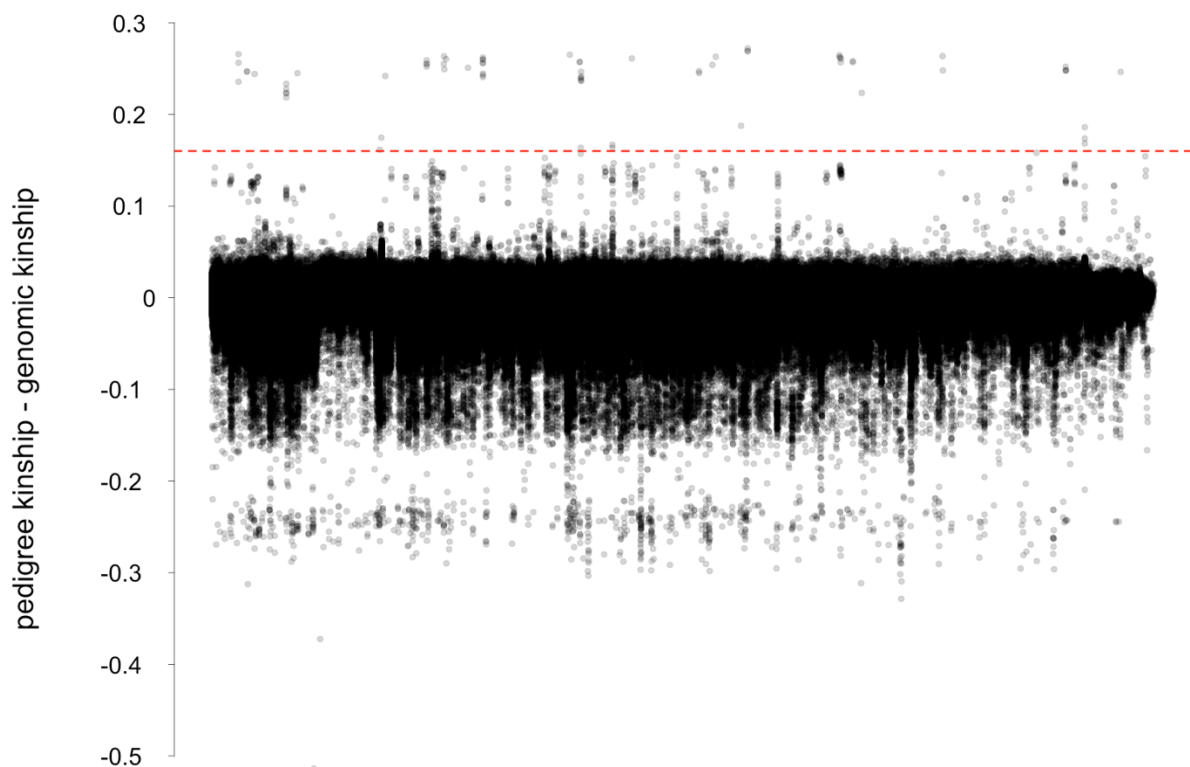


Figure 6: Representation of the difference between the pedigree and genomic kinship. Each point represents a pair of individuals.

If the genomic kinship is larger than the pedigree kinship, it indicates that our pedigree is missing links. This is likely to happen since the birds can nest outside our study area. We can't really "fix" these links so we won't focus on these here. What we are really interested

in is when the pedigree kinship is larger than the genomic kinship. This either means that we did not sequence the correct individual or that the pedigree is wrong. We manually investigated the 71 pairs of individuals for which the difference between the pedigree and genomic kinship was larger than 0.16 (the pairs above the dotted red-line in figure 6). These 71 pairs included 93 unique individuals. Manual investigation consisted in checking if the individuals really were who we thought they were by looking at their kinship with other individuals from the same family such as siblings or offspring from different clutches. If we could not confirm the identity of the samples, we removed them from any analyses that required the use of phenotypic data. If we could confirm their identity, we kept the sample. If we found that the pedigree was wrong, we tried to identify the true parents by looking for high kinship with other individuals of the data set.

We could confirm that the individuals we sequenced really were who we thought they were for 76 individuals. We found that the pedigree was wrong in 50% of the cases. For half of these cases, we could fix the pedigree (by identifying the true parent). For 15 samples, we could not confirm that the individuals we sequenced really were who we think they were and had to remove them from any analyses using phenotypic data. We also removed 2 samples because they looked like genomic twins. Finally, we identified a three way sample swap and subsequently corrected the link between genotypic and phenotypic data. Precise list of samples which were removed as well as a detailed description of each pair we investigate can be found in [GitHub](#). Differences between pedigree and genomic kinship after sample removal and pedigree correction are presented in figure 7.

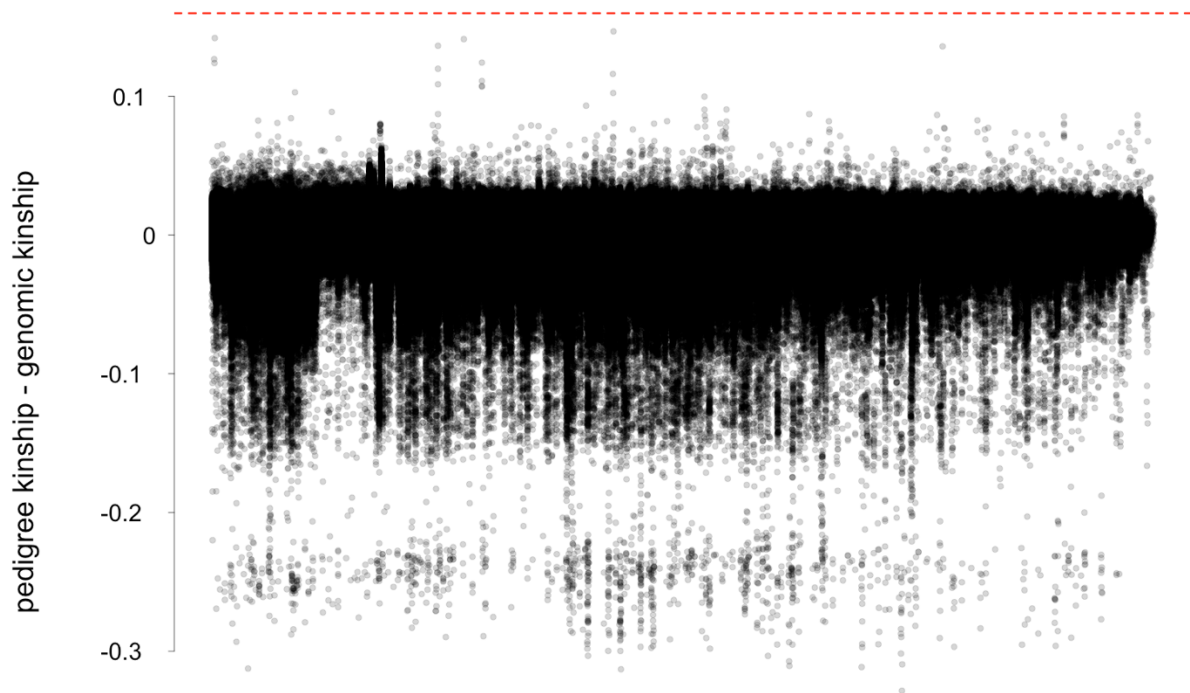


Figure 7: Representation of the difference between the pedigree and genomic kinship after the pedigree correction. Each point represents a pair of individuals.

References

- [1] A.M. Bolger, M. Lohse, and B. Usadel. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15):2114–2120, 2014.
- [2] P. Danecek, J.K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M.O. Pollard, A. Whitwham, T. Keane, S.A. McCarthy, R.M. Davies, and H. Li. Twelve years of *SAMtools* and *BCFtools*. *GigaScience*, 10(2):giab008, 2021.
- [3] J. Goudet. *hierfstat*, a package for r to compute and test hierarchical f-statistics. *Molecular Ecology Notes*, 5(1):184–186, 2005.
- [4] H. Li. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, 27(21):2987–2993, 2011.
- [5] H. Li and R. Durbin. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, 25(14):1754–1760, 2009.
- [6] A.P. Machado, T. Cumer, C. Iseli, E. Beaudoin, A-L. Ducrest, M. Dupasquier, N. Guex, K. Dichmann, R. Lourenço, J. Lusby, H-D. Martens, L. Prévost, D. Ramsden, A. Roulin, and J. Goudet. Unexpected post-glacial colonisation route explains the white colour of barn owls (*tyto alba*) from the british isles. *Molecular Ecology*, 31(2):482–497, 2022.
- [7] S. Rubinacci, D.M. Ribeiro, R.J. Hofmeister, and O. Delaneau. Efficient phasing and imputation of low-coverage sequencing data using large reference panels. *Nature Genetics*, 53(1):120–126, 2021.
- [8] J.P. Sinnwell, T.M. Therneau, and D.J. Schaid. The kinship2 R Package for Pedigree Data. *Human Heredity*, 78(2):91–93, 2014.
- [9] D. Speed and D.J. Balding. Relatedness in the post-genomic era: is it still useful? *Nature reviews. Genetics*, 16(1):33–44, 2015.

ANNEX IV

Authors : Tristan Cumer^{1,2}, Eléonore Lavanchy^{1,2}, Alexandros Topaloudis^{1,2}, Jérôme Goudet^{1,2}

¹ Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland

² Swiss Institute of Bioinformatics, University of Lausanne, Lausanne, Switzerland

This annex describes all the steps implicated in the genomic data generation related to the sexual chromosome (the Z chromosome) for both high (reference panel) and low coverage data sets. It includes the description of identification, genotyping, phasing (at both high and low coverage) and filtering of recombinant regions of the sexual chromosome.

Authors contributions: TC identified the recombinant regions of the sexual chromosome. EL performed the genotyping (for both low and high coverage data sets) and the (read-based) phasing of the low coverage samples. TC performed the phasing of the high coverage samples. AT performed the filtering.

Identification of the PseudoAutosomal Region

In order to call the genotypes of the individuals in the Z chromosome, we first identified the pseudoautosomal region and the non-pseudoautosomal regions (PAR and non-PAR respectively). We extracted the coverage of each base pair for each individual along the Z chromosome (Super-Scaffold 13 and Super-Scaffold 42, [4]) using `samtools depth` (with the `-Q 60` and `-aa` option) [3]. This coverage was then averaged in 5 kb windows and normalised by dividing the coverage of each window with the genome-wide average coverage of this individual (see Annex III for the calculation of the mean coverage of each individual). Finally, we computed a mean normalised coverage across all the individuals of each sex in each window. We then classified the windows as non-PAR or PAR based on their mean normalised coverage in males and females (figure 1); First, we discarded windows with a mean normalised coverage lower than 0.1 in females and 0.2 in males, or a mean normalised coverage higher than two in at least one sex. Then we classified as non-PAR windows the ones with a female coverage 1.05 times (i.e. 5%) lower than in males (in orange in figure 1). PAR windows were identified as windows with a coverage in females similar to the coverage in males (in red in figure 1).

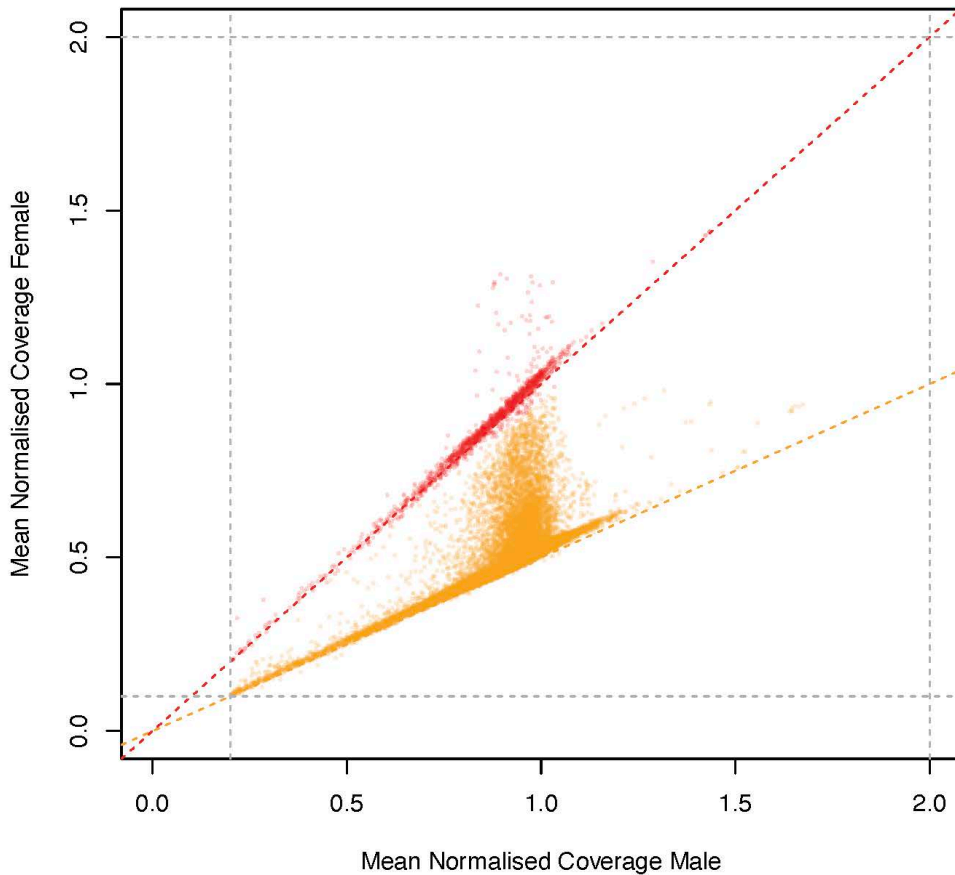


Figure 1: Mean normalised coverage in males and females for each 5kb window along the Z chromosome (Super-Scaffolds 13 and 42). Each dot represents one 5kb window. Red dots depict windows classified as PAR since the coverage was similar in males and females. Orange windows depict windows classified as non-PAR since the coverage was different between males and females. Windows with a mean normalised coverage lower than 0.1 in females and 0.2 in males, or a mean normalised coverage higher than two in at least one sexes are not represented in this figure.

We finally identified PAR and non-PAR regions along the Z chromosome (Super-Scaffolds 13 and 42) by merging adjacent windows into larger regions. Windows with a bad coverage surrounded by the same type of regions (either PAR or non-PAR) were included in a final merging step. Results of this classification are presented in table 1.

Table 1: Coordinate of PAR and non-PAR regions along the Super-Scaffolds 13 and 42.

Chromosome	Start	Stop	PAR/non-PAR
Super-Scaffold_13	1	48838493	non-PAR
Super-Scaffold_42	1	4415000	PAR
Super-Scaffold_42	4415001	7364000	non-PAR
Super-Scaffold_42	7364001	7389000	PAR
Super-Scaffold_42	7389001	41426968	non-PAR

SNPs calling on the Z for the Reference Panel

For the scaffolds belonging to chromosome Z (Super-Scaffold 3 and Super-Scaffold 42), genotyping of samples was performed separately for PAR and non-PAR regions. PAR regions were called and phased as for the rest of the genome since both sexes are diploid (Annex I and II). The non-PAR region was called separately for male and female samples since there is a difference in ploidy. For female birds *GATK's Haplotypecaller v4.2.6* was run with a ploidy of 1, signifying haploid samples. Joint genotyping was performed separately in the two sexes for the non-PAR region using *GATK's GenotypeGVCFs [5]* in accordance with the rest of the genome. Before merging the individuals of the two sexes, females were transformed to diploids by doubling the haploid genotypes.

Quality control filters ($QD < 2.0$, $QUAL < 30$, $SOR > 3.0$, $FS > 60.0$, $MQ < 40.0$, $MQRankSum < -12.5$ and $ReadPosRankSum < -8.0$) were applied as for the rest of the genome along with the mappability mask (see Annex I). For the individual depth filter we filtered males as for the rest of the genome (minimum 5, maximum mean + 3 x SD - see Annex I) but modified female depth filters on the non-PAR region due to the different ploidy. Thus for the haploid female regions we set the minimum read depth to 2 and recalculated the mean and SD of the depth distribution using all SNPs in Super-Scaffold 13 since as mentioned above all of this scaffold belongs to the non-PAR region and is thus haploid in females. A description of individual depth filters is given below (table 2).

Table 2: Individual depth filtering guide

Sex/Region	Ploidy	Minimum depth	Maximum depth	Regions used to calculate mean and SD
All/Autosomal genome	2	5	Mean + 3 SD	Super-Scaffold_1, Super-Scaffold_2
Male/Z chr	2	5	Mean + 3 SD	Super-Scaffold_1, Super-Scaffold_2
Female/Z chr PAR	2	5	Mean + 3 SD	Super-Scaffold_1, Super-Scaffold_2
Female/Z chr nonPAR	1	2	Mean + 3 SD	Super-Scaffold_13

SNPs calling on the Z for the Low coverage samples

Sequencing, generation of BAM files, and extraction of variant positions for the Z chromosome were conducted as described in Annex III for the rest of the genome. Similarly to high coverage samples, PAR regions were genotyped and phased as autosomes since both sexes are diploid (described in Annex III).

We followed the [GLIMPSE website tutorial for the X chromosome in humans](#) for genotyping, imputation and phasing of the non-PAR regions. Non-PAR regions were genotyped with `BCFTOOLS` v.1.15.1 [1] `mpileup` and `call` methods [2] with ploidy 1 for females and 2 for males (defined with the `--ploidy-file` parameter). Variants were only called at variant positions in the reference panel (`-T` option) and for alleles present in the same panel (option `-C allele`). Imputation and phasing were performed with `GLIMPSE` v1.1.1. Chunking is the first step of the `GLIMPSE` pipeline and involves splitting the chromosomes into (overlapping) chunks. The size of these chunks is automatically estimated based on several parameters, including the amount of missing data and the length of the chromosome. This is implemented in the `GLIMPSE_chunk` method. We performed chunking exactly as described for the autosomes (in Annex III): with default parameters (`-window-size 2000000` and `--buffer-size 200000`). For the imputation and phasing, we used `GLIMPSE_phase` with a burn-in of 100 (`--burnin 100`) and 15 main iterations (`--main 15`). We set the effective population size at 10,000 individuals (`--ne 10000`) and the ploidy at 1 for females and 2 for males (defined with the `--samples-file` parameter). The next step is

ligation, which involves merging the different chunks together without losing the phase information. Ligation was performed with the *GLIMPSE_ligate* method for males and females separately: we first split the individuals according to their ploidy and ligated both sexes independently. Finally, we generated haplotype-level genomic information with the *GLIMPSE_sample* method. We identified the most likely haplotype (with the *--solve* option) for males and females separately. We then merged the VCFs of both sexes and manually modified females' non-PAR regions from haploids to homozygous diploids.

References

- [1] P. Danecek, J.K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M.O. Pollard, A. Whitwham, T. Keane, S.A. McCarthy, R.M. Davies, and H. Li. Twelve years of *SAMtools* and *BCFtools*. *GigaScience*, 10(2):giab008, 2021.
- [2] H. Li. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, 27(21):2987–2993, 2011.
- [3] H. Li, B. Handsaker, A. Wysoker, T.Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and *SAMtools*. *Bioinformatics*, 25(16):2078–2079, 2009.
- [4] A.P. Machado, T. Cumer, C. Iseli, E. Beaudoin, A-L. Ducrest, M. Dupasquier, N. Guex, K. Dichmann, R. Lourenço, J. Lusby, H-D. Martens, L. Prévost, D. Ramsden, A. Roulin, and J. Goudet. Unexpected post-glacial colonisation route explains the white colour of barn owls (*tyto alba*) from the british isles. *Molecular Ecology*, 31(2):482–497, 2022.
- [5] R. Poplin, V. Ruano-Rubio, M.A. DePristo, T.J. Fennell, M.O. Carneiro, G.A. Van Der Auwera, D.E. Kling, L.D. Gauthier, A. Levy-Moonshine, D. Roazen, K. Shakir, J. Thibault, S. Chandran, C. Whelan, M. Lek, S. Gabriel, M.J. Daly, B. Neale, D.G. MacArthur, and E. Banks. Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv*, 2017.