Conservation Genetics Resources
https://doi.org/10.1007/s12686-024-01369-z

METHODS AND RESOURCES

Pure wild forest reindeer (*Rangifer tarandus fennicus*) or hybrids? A whole-genome sequencing approach to solve the taxonomical status

Melak Weldenegodguad[1] · Milla Niemi[2] · Sakari Mykrä-Pohja[3] · Kisun Pokharel[4] · Tuula-Marjatta Hamama[4] · Antti Paasivaara[5] · Juha Kantanen[4]

Received: 17 August 2023 / Accepted: 2 October 2024

Abstract
In Finland, the geographic distribution of domestic reindeer (*Rangifer tarandus tarandus*) and Finnish wild forest reindeer (*Rangifer tarandus fennicus*) partly overlap in the vicinity of the southern border of the reindeer herding area. Additionally, domestic reindeer are occasionally kept as pets within the distribution range of the wild forest reindeer. Hybridisation of these two subspecies is one of the major threats for the wild forest reindeer population. Concerns about potential hybridisation served as the catalyst also for this study, which we aimed to clarify the taxonomic status of presumed wild forest reindeer individuals intended as founder individuals for a reintroduction project. To do this, we resequenced genomes of four *Rangifer tarandus* individuals with unknown taxonomical status and investigated their ancestries by comparing the genomic data with the existing resequenced data of the Finnish domestic reindeer and Finnish wild forest reindeer. The genetic relationship investigations indicated that all individuals we analysed were pure wild forest reindeer, making them suitable as founder individuals for the reintroduction project. Thus, our study provided critical knowledge for practical conservation action, where it was essential to recognise each individual's origin. In the future, it will also offer novel insights into the spread of native wild forest reindeer to new geographic regions in Finland. For subsequent studies, additional resequenced genomic data of *Rangifer* individuals will be needed to develop an ancestry information marker panel of single nucleotide polymorphisms for rapid and cost-effective identification of hybrid individuals of these subspecies.

Keywords Admixture · Wild forest reindeer · Genome · Hybridisation · Reintroduction · Domestic reindeer

Introduction

Hybridisation or crossbreeding between individuals of different species may occur naturally in 10% of animal species (Mallet 2005). While it could act as a source of genetic variation, especially hybridisation between wild populations and their domesticated or captive-reproduced relatives is an increasing conservation concern worldwide (e.g., Randi 2008; Stroupe et al. 2022; Alice Brambilla et al. 2023; Šprem et al. 2023). The introduction of non-native species to new areas has led to hybridisation or crossbreeding, as with several European ungulate populations, for example (e.g. Randi 2005; Iacolina et al. 2019).

The forest-dwelling Finnish wild forest reindeer (*Rangifer tarandus fennicus*), a native subspecies of reindeer with circumpolar distribution, nowadays exists only in Finland, Russian Karelia and the westernmost part of the Arkhangelsk oblast (Panchenko et al. 2021). Based on aerial censuses, there were about 3,000 wild forest reindeer

✉ Milla Niemi
milla.niemi@metsa.fi

✉ Juha Kantanen
juha.kantanen@luke.fi

1 Natural Resources Institute Finland (Luke), Latokartanonkaari 9, Helsinki FI-00790, Finland

2 Wildlife Service Finland, Sairaalantie 3b, Ivalo FI-99801, Finland

3 Wildlife Service Finland, Pohjoispuisto 7, Pori FI-28100, Finland

4 Natural Resources Institute Finland (Luke), Tietotie 4, Jokioinen FI-31600, Finland

5 Natural Resources Institute Finland (Luke), Paavo Havaksentie 3, Oulu FI-90570, Finland

Published online: 12 October 2024

in the late winter of 2023 in Finland (Natural Resources Institute Finland, unpublished data), divided over the two main subpopulations, Kainuu and Suomenselkä (Paasivaara et al. 2021, 2022). The Suomenselkä subpopulation was introduced from Kainuu in the early 1980s, and there is no confirmed gene flow between subpopulations. In addition, two new subpopulations have been reintroduced recently (Mykrä-Pohja and Niemi 2022). The Russian population has decreased during the last few decades, the last estimate being approximately 2,300 individuals (Panchenko et al. 2017; Danilov et al. 2020), leading to the global population of about 5,000–6,000 individuals. In contrast to decreasing wild reindeer populations, Klokov (2007) estimated that there are nearly 2,000,000 domestic reindeer (*Rangifer tarandus*) globally, of which one-third are distributed in Fennoscandia. There, the maximum number of domestic reindeer is regulated through legislation, leading somewhat stable population sizes about 200 000 individuals in Finland.

The recent whole-genome sequencing study evidenced that the wild forest reindeer are genetically distinct from the domestic reindeer (Pokharel et al. 2023). Moreover, the archaeo-osteological study by (Rankama and Ukkonen 2001) indicated separate origins for northern Fennoscandian forest and tundra reindeer; the forest reindeer may have originally spread to Finland from the east circa 7,500 years ago, following the retracting ice margin, while the tundra reindeer may have colonised northern Fennoscandia especially along the Norwegian narrow ice-free coastal zone.

Although the Fennoscandian domestic reindeer has been domesticated from the tundra reindeer (also called the mountain reindeer, *Rangifer tarandus tarandus*) (Røed et al. 2021), the domestic reindeer and the wild forest reindeer partly have a common history as suggested by Heino et al. (2021). They studied ancient reindeer bone samples from indigenous Sámi offering sites and found that mitochondrial DNA (mtDNA) haplotypes common to wild forest reindeer were present even in the most northern part of present-day Finland, that is, in the current reindeer herding area. They began to be replaced by mtDNA haplotypes found in domestic reindeer starting between 1400 and 1600 CE, indicating the growing importance of reindeer herding during those centuries. Correspondingly, the wild forest reindeer were hunted in a reindeer herding area, and the last herd was not shot until 1883 (Montonen 1974).

Due to their long history, domestic reindeer and wild forest reindeer are capable to mate and produce fertile offspring. Crossbreeding has been recognised as a major threat for the wild forest reindeer genome (Härkönen and Bisi 2007), and an active prevention program targets preventing hybridisation and removing possible hybrids from the Finnish subpopulations of wild forest reindeer (Niemi et al. 2021). Most encounters leading to crossbreeding between the domestic reindeer and wild forest reindeer occur near the southern border of the reindeer herding area, where about 13 000 domestic reindeers exist permanently in six herding districts. However, there is sporadic, small-scale reindeer keeping also in the southern part of Finland, meaning that hybridisation could occur there as well if animals escape from their owner (Wildlife Service Finland, unpublished data). Even if the appearance of hybrids in the wild forest reindeer population seems to be low (Pokharel et al. 2023), the expanding distribution area of wild forest reindeer will most likely lead to an increased number of encounters with domesticated and wild reindeer, underlining the need for tools to test a possible hybridisation.

In this study, we analysed the genome of four supposed wild forest reindeer individuals to identify their possible hybridisation status and to recognise which subpopulation the individuals originated. Three of these animals were intended to be used as founder individuals in a wild forest reindeer reintroduction project, underlining the importance of their genetic origin. For determination of their taxonomical status (and possible hybridisation), we used the existing whole-genome sequencing data of the Finnish domestic reindeer and wild forest reindeer (Weldenegodguad et al. 2020; Pokharel et al. 2023).

Table 1 Summary of samples used in whole-genome sequence analysis. The four *Rangifer* individuals studied were investigated during this study, whereas other samples had been sequenced earlier and served as references

Population	No. of individuals
Finnish domestic reindeer (*Rangifer tarandus tarandus*)	10
Finnish wild forest reindeer (*Rangifer tarandus fennicus*)	6
Studied *Rangifer tarandus* individuals 1, 2, 3, 4	4
Alaskan wild caribou (*Rangifer tarandus caribou*)	1

Materials and methods

Collecting samples

In the winter of 2019, a wild forest reindeer female and its female offspring (*individual 1*; see Table 1 and Supplementary Tables S1 and S2) were detected in the reindeer herding area in Kainuu, where a 90-kilometre-long fence has been built to separate domestic reindeer and wild forest reindeer populations (Niemi et al. 2021). Both individuals were caught (Permit criteria of the Finnish National Animal Experiment Board; Decisions on the granting of license of team Regional State Administrative Agency for Southern Finland, ESAVI: 6336/04.10.03/2012, 587/04.10.07/2016 and 23666/2018) and transported to the breeding enclosure located in Lauhanvuori National Park (62.1521° N, 22.1751°

E), which is one of two reintroduction sites of wild forest reindeer in the EU-funded project WildForestReindeerLIFE (Mykrä-Pohja and Niemi 2022). Because those transported individuals were intended to act as founder individuals for the newly reintroduced wild forest reindeer population and animals were found from the reindeer herding area, the calf's possible hybridisation status needed to be tested. As the adult female was pregnant and gave a birth in the enclosure in the spring of 2019, the status of that newborn male offspring (*individual 2*) also had to be checked. The young female was anaesthetised and sampled in Autumn 2020 and the male in the summer of 2021.

In January 2022, the police captured a young *Rangifer* male (*individual 3*) from Helsinki (YLE 2022), which is located more than 200 km to the south of the wild forest reindeer reintroduction areas and 300 km from the distribution area of the closest wild population (Fig. 1) (Pöllänen et al. 2023). After recovering in the Helsinki Zoo, that individual was also transported to the wild forest reindeer breeding enclosure in Lauhanvuori National Park. Its mtDNA was analysed before transportation (Korkeasaari Zoo 2022), but as the analysis did not reveal possible crossbreeding with a wild forest female and domestic reindeer male (Amorim et al. 2020), the possible hybridisation status needed to be confirmed through this study.

With these three samples related to the reintroduction of wild forest reindeer, an older sample (*individual 4*) from a supposed wild forest reindeer was also analysed. This individual was found dead in Joutsa, Finland, in the autumn of 2017 outside the known distribution area of wild forest

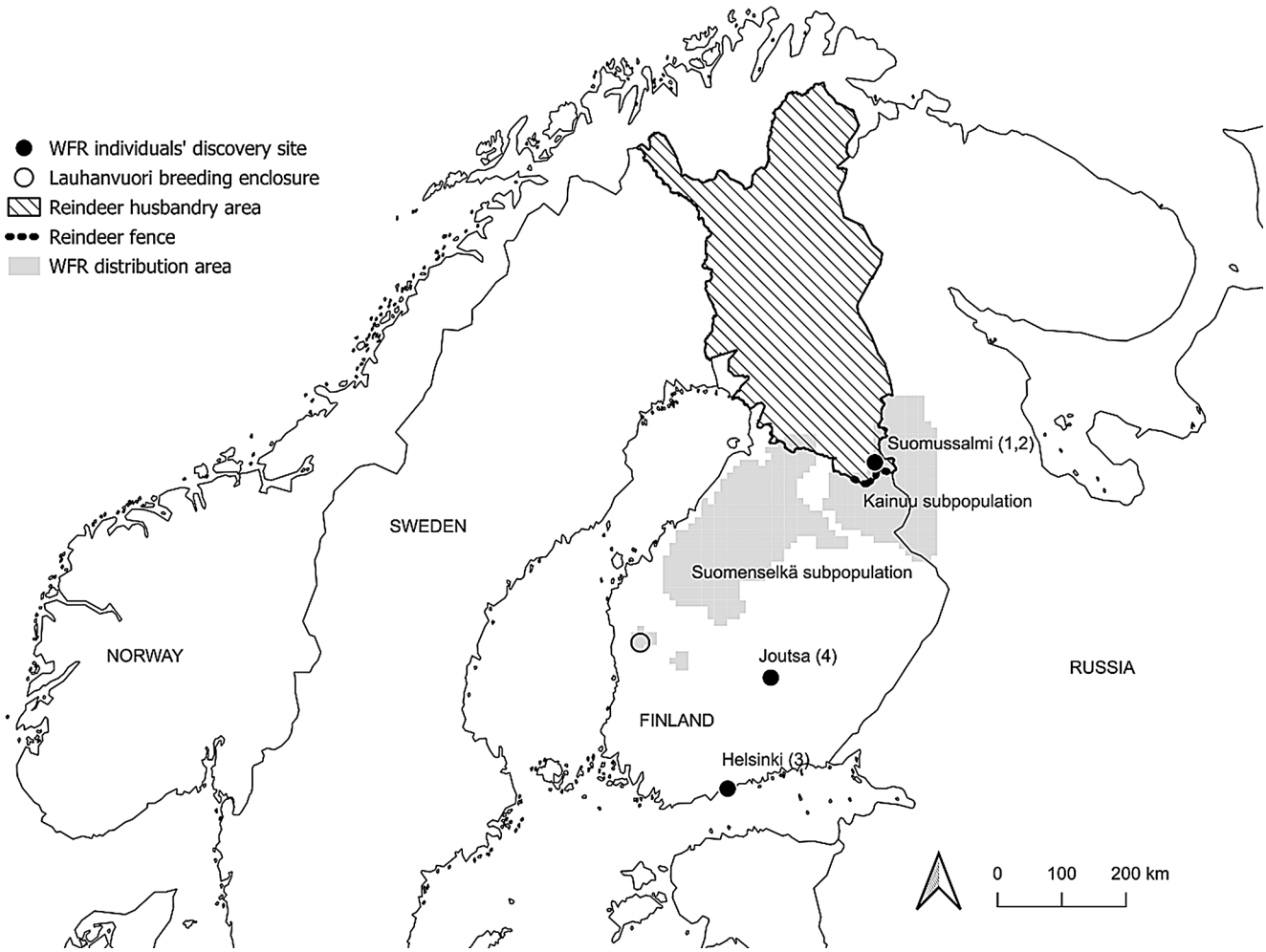

Fig. 1 Discovery sites of studied wild forest reindeer individuals and breeding enclosure in Lauhanvuori National Park. The subpopulations of wild forest reindeer (WFR) are marked with grey and reindeer herding (or reindeer husbandry) area with diagonal lines. This information was provided by the Natural Resources Institute Finland (wild forest reindeer distribution), the Reindeer Herders' Association (reindeer herding area) and Eurostat (country borders). The abbreviation 'WFR' refers to 'wild forest reindeer'. The fence is located only in the Kuhmo region in East Finland. On the map, Suomussalmi (1, 2) refers to the geographical origin of *individual 1* and *individual 2* in Kainuu region, while *individual 3* was found in Helsinki and individual 4 in Joutsa. In Finland, reindeer herding area (or reindeer husbandry area) covers Lapland region and northern parts of North Ostrobothnia and Kainuu region. The reindeer herding area covers 36% of Finland's total area (122,936 sq. kilometers)

reindeer (Fig. 1). DNA was extracted from blood (*individuals 1*, *2* and *3*) and from hair samples (*individual 4*).

Whole-genome sequencing

To examine the genetic relationships between the Finnish domestic and the four presumed wild forest reindeer individuals described above, we used a total of 21 resequenced reindeer samples (Table 1 and Supplementary Table S1). The previous data included genome sequences of 10 Finnish domestic reindeer, six Finnish wild forest reindeer and, as an outgroup, one wild caribou from Alaska, USA (Weldenegodguad et al. 2020; Pokharel et al. 2023); these data were generated at Beijing Genomics Institute using the Illumina HiSeq 4000 platform. The new samples described above and investigated here were sequenced using the Illumina HiSeq 4000 with a 150 bp paired-end strategy at Novogene.

Bioinformatic analysis

Prior to the bioinformatic analysis, the quality of the raw reads was assessed using FastQC software v0.11.8 (Andrews 2010). MultiQC v1.9 (Ewels et al. 2016) was used to summarise the quality control reports of all samples from FastQC. High-quality clean reads for each sample were mapped against the domestic reindeer reference genome (Weldenegodguad et al. 2020) using BWA v0.7.17 (Li and Durbin 2009) with default parameters. Following mapping, the SAM files generated by BWA were converted to binary equivalent BAM files and sorted using Picard tools v2.21.4 (https://broadinstitute.github.io/picard/). PCR duplicates from the aligned reads were removed using Picard tools.

After preprocessing the mapped reads, the Genome Analysis Toolkit (GATK) best practices pipeline (Van der Auwera et al. 2013) was used to identify high-quality variants (single nucleotide polymorphisms [SNPs] and indels) from the uniquely mapped reads. Variants were called using the GATK v4.2.0.0 HaplotypeCaller from the uniquely mapped reads. Based on GATK4 user guide recommendations, sequencing and alignment artifacts from the variants were discarded using the following parameters: FisherStrand (FS) > 60.0, RMS Mapping Quality (MQ) < 40.0, Mapping Quality Rank Sum Test (MQRankSum) < − 8.0, QualByDepth (QD) < 2.0, ReadPosRankSum < − 8.0 and StrandOddsRatio (SOR) > 3.0 for SNPs and FS > 200.0, MQ < 40.0, QD < 2.0, ReadPosRankSum < − 8.0 and SOR > 5.0 for indels.

To examine the genetic relatedness between the samples, a principal component analysis (PCA) was conducted using the identified SNPs with a Bioconductor R package SNPRelate, v1.28.0 (Zheng et al. 2012). SNPs were first filtered based on linkage disequilibrium (LD) using the snpgdsLDpruning function in the SNPRelate package to avoid strong influence of linked SNP clusters (see Pokharel et al. 2023). We used 0.2 LD thresholds for filtering the SNPs. The R function snpgdsPCA was used to perform a PCA plot.

Moreover, the identified SNPs were utilised to generate a Neighbour-Joining phylogenetic tree using SNPhylo v. 20,160,204 (Lee et al. 2014). Before constructing the tree, SNPs were further filtered using SNPhylo with the following filter parameters: minimum depth of coverage > 5, percentage of low-coverage samples < 5%, percentage of samples with no SNP information < 5%, LD < 0.1 and minor allele frequency > 0.05. The filtered SNPs were then concatenated to generate sequences and used to perform multiple alignments using MUSCLE v3.8.31 (Edgar 2004). A Neighbour-Joining tree was then constructed by running DNAML programs in the PHYLIP package (Felsenstein 2005). Bootstrap analysis was performed using the "phangorn" package with 100 replications (Guindon et al. 2010). FigTree v1.4.4 (Rambaut 2018) was used to visualise the phylogenetic tree.

Furthermore, population admixture based on the identified SNPs was examined using ADMIXTURE software, v.1.3.0 (Alexander et al. 2009). The SNPs in VCF format were first converted into binary PLINK format for ADMIXTURE input using PLINK v1.90b6.21 (Purcell et al. 2007). ADMIXTURE was then examined using different ancestral clusters (K) ranging from two to five. The present *R. tarandus* populations (domestic reindeer, wild forest reindeer and caribou) grouped into three different clusters in the previous phylogenetic study (Pokharel et al. 2023). To estimate standard errors, the bootstrap parameter was set to 100 replicates.

Results

Whole-genome sequence data

In this study, a total of 441 gigabases (Gb) of paired-end whole-genome sequence data were generated from 21 samples (Table 1 and Supplementary Table S2) to examine the genetic relatedness between the samples. The quality of all the samples was found to be high and, hence, quality filtering was not performed. On average, each sample had 274 million (M) and 41.18 Gb clean reads and bases, respectively (Supplementary Table S2). The reads were successfully aligned to the assembled reference genome, with an average alignment rate of 98.31%, and represented 10-fold coverage (Supplementary Table S2).

Variant calling

A total of 28.1 M SNPs were detected using the uniquely mapped reads across all 21 samples (Supplementary Table S3). The average number of SNPs detected per individual was 8.1 M. In the domestic reindeer, the average number of SNPs was 7.9 M, in the wild forest reindeer it was slightly higher at 8.2 M and in the four samples investigated here it was 8.5 M. On average, the calculated transition-to-transversion (Ts/Tv) ratios were 1.90 per individual (Supplementary Table S3). The observed Ts/Tv ratios were found to be slightly lower than that of humans (2.1) (Lachance et al. 2012) and bovines (2.2) (Stothard et al. 2011; Choi et al. 2014, 2015). Moreover, a total of 3.9 M indels were detected across all 21 samples, and on average 1.23 M indels were detected per individual.

Population structure analysis

PCA and phylogenetic tree analysis were used to examine the genetic relationships between all 21 animals. In the phylogenetic tree, the wild caribou from Alaska was an outgroup. The PCA plot and phylogenetic tree revealed two major groups: the Finnish domestic reindeer versus the Finnish wild forest reindeer (Figs. 2 and 3). Both the PCA plot and phylogenetic tree clearly indicated that the two animals captured from the reindeer herding area (*individual 1* and *2*), the *Rangifer* sp. individuals found in Helsinki (*individual 3*) and Joutsa (*individual 4* which was dead), were pure *Rangifer tarandus fennicus*, and no sign of hybridisation with domestic reindeer was detected.

Results from analysis of population structure and admixture are seen in Fig. 4 showing the obtained structure plots of the four populations for all the K values. Cross-validation errors of the cluster numbers were lowest for K values two and three, increasing after that as K increased (Supplementary Table S4). This suggests that either two or three would be the optimal number of ancestral populations. As shown in Fig. 4, at K = 3, individuals 1, 2, 3 and 4 showed ancestries with Finnish forest reindeer populations. This clustering is consistent with the results of PCA and phylogenetic tree (Figs. 2 and 3). Moreover, our analyses (Figs. 3 and 4) clearly show that the individuals 1 and 2 grouped with the forest reindeer samples which were collected in the Kainuu subpopulation (RTF17, RTF31, RTF84, RTF96) whereas the individuals 3 and 4 clustered with the samples RTF6 and RTF7 collected in Suomenselkä subpopulation.

Discussion

Individuals' taxonomical status and hybridisation

We sequenced the whole genome of four presumed wild forest reindeer individuals to confirm their taxonomical status and possible hybridisation. Our results showed that all four individuals tested were wild forest reindeer, and no marks of hybridisation were detected. Thus, three of these wild forest reindeer, which were kept in the breeding enclosure in Lauhanvuori National Park, were considered suitable as founder individuals for reintroduced subpopulation. Two tested males were then released into the Lauhanvuori National Park, but unfortunately, the female died in the enclosure.

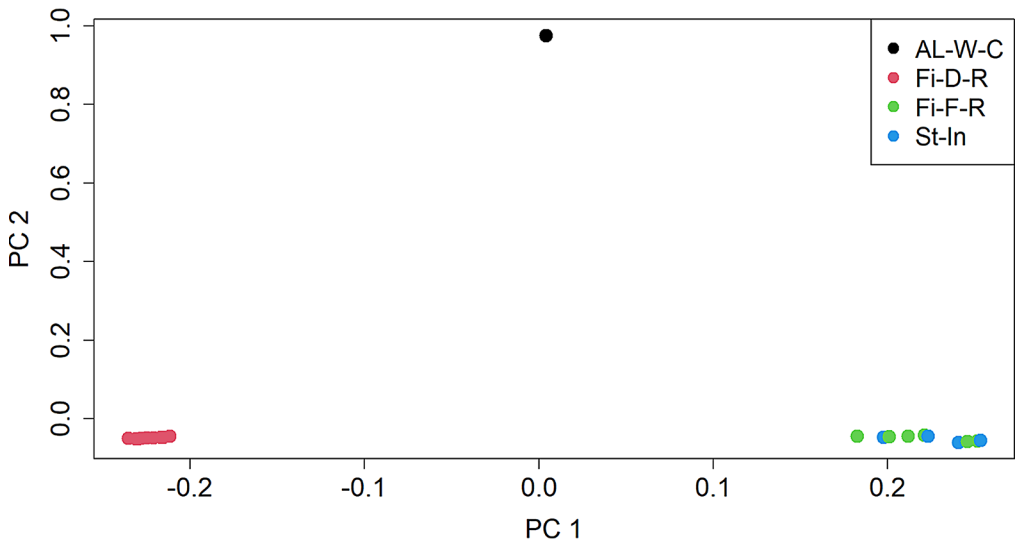

Fig. 2 PCA plot for 21 resequenced animals. Finnish domestic reindeer (Fi-D-R) clustered separately from the Finnish wild forest (Fi-F-R) individuals. Four new samples included in this study (St-In) were grouped with Fi-F-R individuals. A sample representing Alaskan wild caribou (AL-W-C) stood out as an outgroup

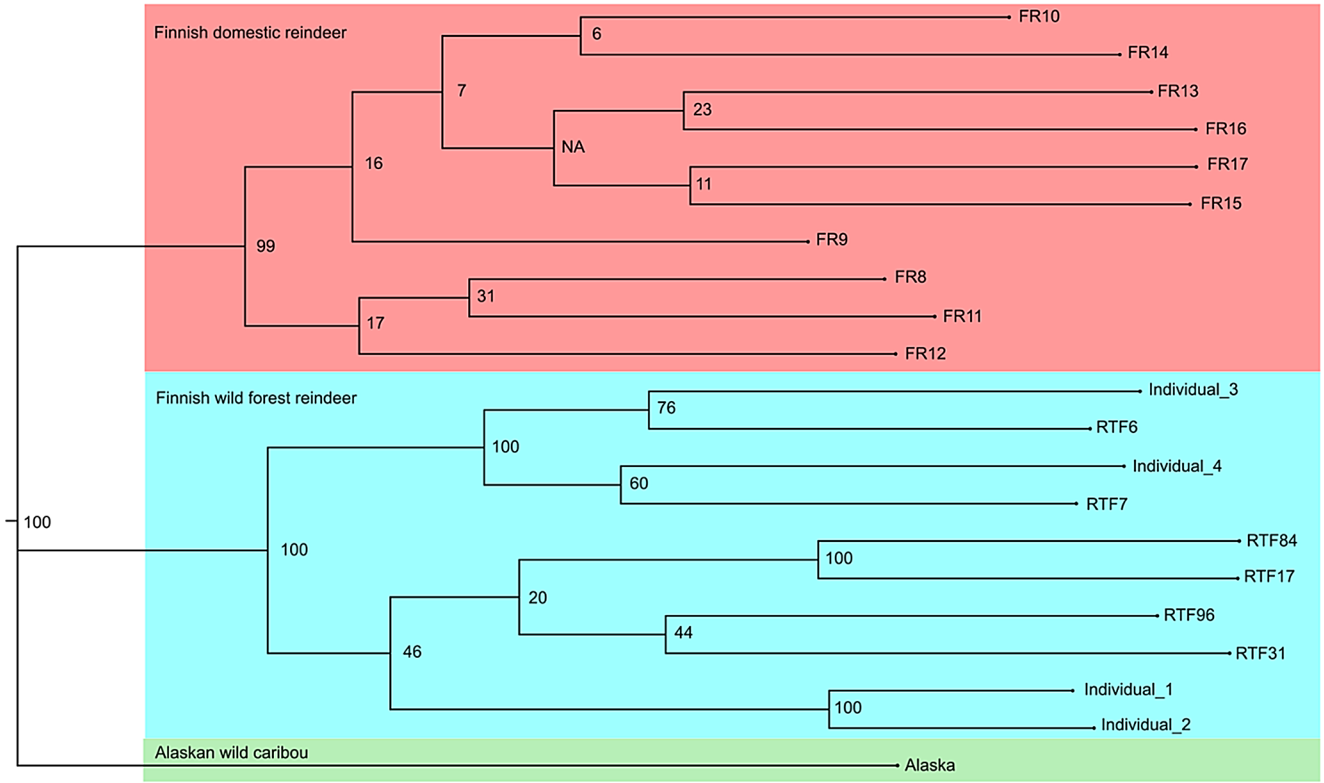

Fig. 3 Phylogenetic analysis indicated two main clusters of reindeer populations. (1) Finnish domestic reindeer (FR-individuals) and (2) Finnish wild forest reindeer (RTF-individuals)/studied individuals. Bootstrap confidence values are shown at each branch

We selected to use whole-genome sequencing data in our study because its ability to distinguish wild forest reindeer from domestic reindeer is clearly more powerful than autosomal microsatellites (Weldenegodguad et al. 2020; Pokharel et al. 2023), which have previously been used to examine genetic diversity of wild forest reindeer (e.g. Røed et al. 2008). However, there is a need for a more cost-effective and rapid method to identify potential hybrids. By using so called an ancestry information marker (AIM) approach for whole-genome sequence data analysis, it will be possible to develop an AIM single nucleotide polymorphism (SNP) panel with a limited number SNPs for hybrid identification (see e.g. Somenzi et al. 2020). Although more genomes of both domestic reindeer and wild forest reindeer individuals need to be sequenced, the samples we analysed in this study could be utilized in this developing work. When usable, the SNP-panel would facilitate the routine identification of hybrids when collecting wild forest reindeer or their gametes from the wild. The AIM SNP panel can also be used for example to reveal the poaching of wild forest reindeer in the case that only a skinned carcass or meat is found.

Wildlife Service Finland is currently planning a new reintroduction, where several wild forest reindeer individuals intended as founder animals will be caught and transported from the nature into breeding enclosures. In theory, hybridisation could occur anywhere within the range of the wild forest reindeer, and hybrids with domestic reindeer cannot be always identified based solely on their phenotypes. Therefore, we recommend that any individuals captured from the wild for reintroduction purposes will be tested to detect potential hybridisation to avoid the situation where a new population was established by using individuals carrying genes from a domestic subspecies. However, the use of genetic data offers even a wider range of opportunities to select most suitable founder individuals for reintroductions. Genomic data has recently used for example to identify appropriate founders for woodland caribou conservation breeding program in Jasper National Park in Canada (Cavedon et al. 2023).

The importance of assisted reproduction in conservation purposes has been increasing (e.g. Acevedo and Barfield 2023). In the future, it may be possible to collect semen from wild forest reindeer males killed for example in traffic and utilize them to improve the genetic diversity of the zoo population through assisted reproduction. Such as living founder individuals, the potential hybrid status of these donor animals should be routinely checked.

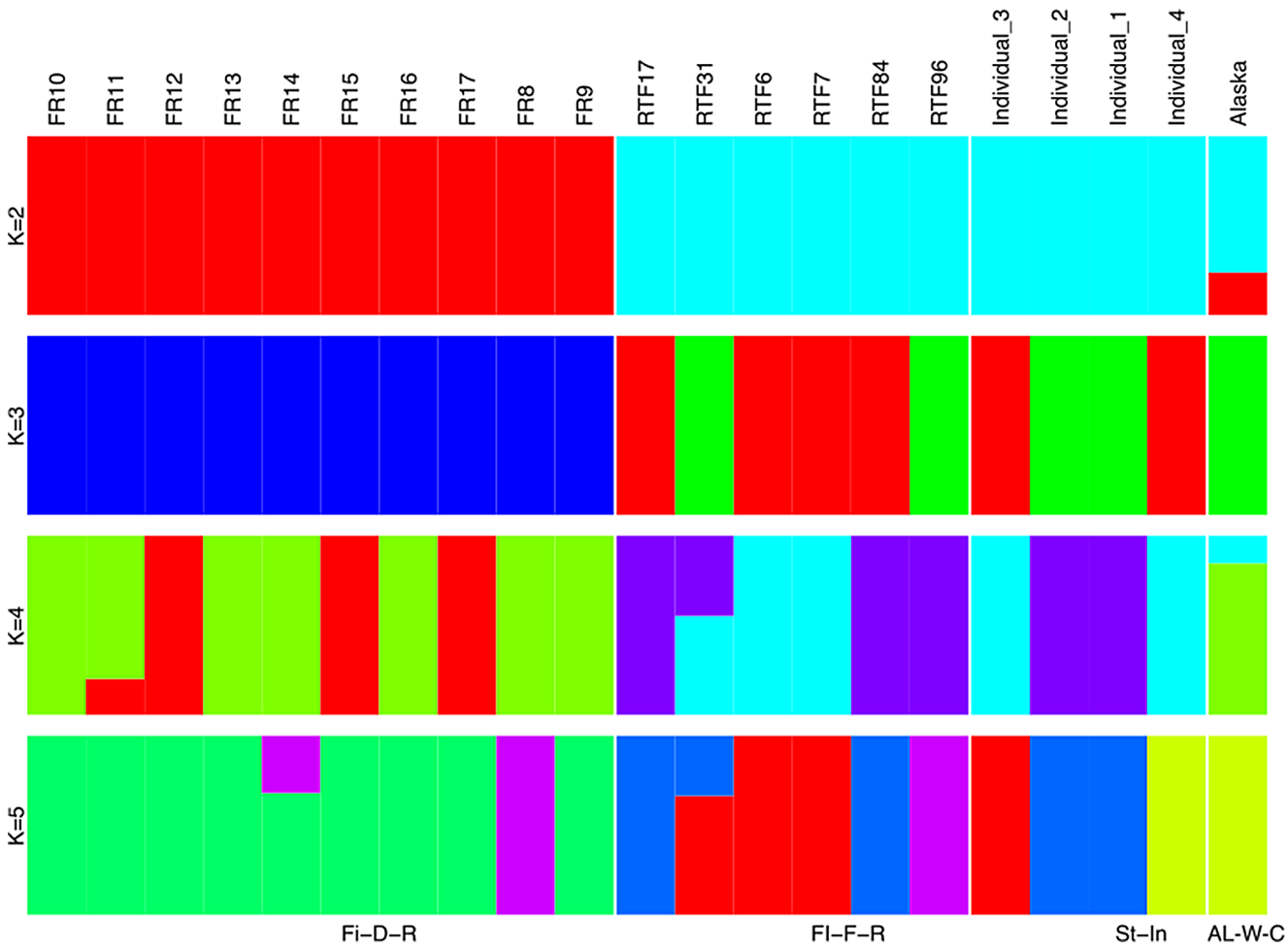

Fig. 4 Population structure analysis of the four populations using ADMIXTURE. (Fi-D-R, Finnish domestic reindeer; Fi-F-R Finnish wild forest reindeer; St-In, Studied individuals, AL-W-C, Alaskan wild caribou). The analysis was repeated with different assumed numbers of clusters (K)

Individuals' origin

All of four individuals we tested displayed high levels of genetic variation although they originated from different subpopulations. Our results indicate that the tested female (*individual 1*) and the young male (*individual 2*) which were born to the adult female captured from the reindeer herding area, originated from the Kainuu subpopulation. The result was highly expected as the pregnant female was caught very close to Kainuu subpopulation. In addition, it had been previously GPS-collared, and its movement history was known from 2011 to 2015 (unpublished data, Natural Resources Institute Finland).

In contrast, the young male (*individual 3*) captured from Helsinki and the one (*individual 4*) that was found after its death in Joutsa, originated from the Suomenselkä subpopulation. Both individuals were far away (about 200 km) from the known areas of this subpopulation. Especially discovery site of *individual 3* was somewhat surprising; calves under the age of one usually follow their mothers and/or the natal herd before first breeding at age two to four (Montonen 1974). The most likely explanation for ending up in the southern coast is that the calf was accidentally lost from its mother and/or normal wintering area, and then wandered following landscape elements until its route was blocked by urban settlement.

The exact age of the adult male (*individual 4*) is not known but it was possibly a young male dispersing from its natal areas. Even adult wild reindeer females typically show a high rate of breeding-site fidelity (Montonen 1974; Schaefer et al. 2000), the natal dispersal patterns are still mainly unknown. However, several observations of young wild forest reindeer individuals, both males and females, have done during the past decades outside their known distribution area (unpublished data, Natural Resources Institute Finland). This is most likely related to the growth and expansion of the Suomenselkä subpopulation (Paasivaara et al. 2021; see also distribution maps in Härkönen and Bisi 2007; Pöllänen et al. 2023). The findings of this study, together with the expanding range of the Suomenselkä population, indicate

that the rate of true dispersal of wild forest reindeer may exist. Potentially, the probability of gene flow between subpopulations may increase accordingly.

Conclusions

In this study, we used the whole-genome sequencing approach to confirm the taxonomical status and possible hybridisation of four *Rangifer* individuals from partially unknown origin. All tested animals were wild forest reindeer, and no marks of hybridisation with domestic reindeer were found. Thus, tested individuals were considered as suitable founders for wild forest reindeer reintroduction project. In addition to their taxonomical status and possible hybridisation, we got information about studied individuals' origin. This helps us to better understand wild forest reindeer natal dispersal, which is still mainly unknown in details. Even relatively rare, hybridisation could occur practically anywhere within the range of the wild forest reindeer, and hybrids cannot be always identified based on their phenotypes. This is why we suggest that each wild forest reindeer captured from the wild for reintroduction purposes or to enhance zoo population, will be tested to detect potential hybridisation. To make this routine, we should develop a SNP panel for cost-effective and faster hybrid identification. For that, whole genome sequencing of more individual is needed. In a long run, the use of genetic methods alongside traditional conservation tools will help us to restrict gene flow from the domestic reindeer to the wild forest reindeer population.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12686-024-01369-z.

Acknowledgements Our special thanks go to veterinarian Heini Niinimäki, who spent several days and nights with MN in the Lauhanvuori breeding enclosure, waiting for the target animals to come close enough to anaesthetize them and collect blood samples. Veterinarian Sanna Sainmaa helped us by taking a blood sample from the individual recovering in Korkeasaari Zoo. In addition, we thank Seppo Salonen and Paula Salonen from Joutsa, who sent a hair sample from a reindeer carcass of unknown origin. Authors also thank Arto Juntunen (deceased), Markku Gavrilov, Marja Hyvärinen and Petri Timonen for their enormous effort catching WFR and collecting wild reindeer blood samples, which were used as references in this study. Mervi Honkatukia, Jaana Peippo, Tiina Reilas and Knut Røed are acknowledged for collaboration to collect the domestic reindeer and caribou samples. The authors wish to acknowledge the CSC-IT Center for Science, Finland, for computational resources.

Author contributions M.N., S.M-P. and J.K. conceived and designed the project. M.N., S.M-P. and A.P. collected the samples of the studied four individuals. T.H. performed the laboratory analyses. M.W. did the bioinformatic analyses. M.W., M.N., S.M-P., A.P., K.P. and J.K. wrote the manuscript. All authors reviewed and approved the final manuscript.

Funding The study was funded by the grant of the project WildForest ReindeerLIFE (LIFE15 NAT/FI/000881) and the grant by the Research Council of Finland in the Arctic Research Programme ARKTIKO (decision number 286040).
Open access funding provided by Natural Resources Institute Finland.

Data availability The whole genome sequence data of the studied individuals (four forest reindeer individuals) is available under ENA (European Nucleotide Archive) study accession code PRJEB66450. The reference sequence genome data is available under ENA study accession: PRJEB65932.

Declarations

Ethical approval Project beneficiaries take full responsibility for the materials included in this publication. Texts and pictures published in this article should not be interpreted to represent the official views or standpoints of the European Commission or the European Union.

Competing interests The authors declare no competing interests.

References

Acevedo C, Barfield JP (2023) Review: Reproductive physiology of bison and application of assisted reproductive technologies to their conservation. animal 17:100842. https://doi.org/10.1016/j.animal.2023.100842

Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. Genome Res 19:1655–1664. https://doi.org/10.1101/gr.094052.109

Alice Brambilla N, Zehnder B, Bassano et al (2023) Genetic evidence of a hybrid swarm between Alpine ibex and domestic goat. bioRxiv 2023.12.22.572967. https://doi.org/10.1101/2023.12.22.572967

Amorim A, Pereira F, Alves C, García O (2020) Species assignment in forensics and the challenge of hybrids. Forensic Sci Int Genet 48:102333. https://doi.org/10.1016/j.fsigen.2020.102333

Andrews S (2010) FASTQC. A quality control tool for high throughput sequence data

Cavedon M, Neufeld L, Finnegan L et al (2023) Genomics of founders for conservation breeding: the Jasper caribou case. Conserv Genet 24:855–867. https://doi.org/10.1007/s10592-023-01540-3

Choi J-W, Liao X, Stothard P et al (2014) Whole-genome analyses of Korean native and holstein cattle breeds by massively parallel sequencing. PLoS ONE 9:e101127. https://doi.org/10.1371/journal.pone.0101127

Choi J-W, Choi B-H, Lee S-H et al (2015) Whole-genome resequencing analysis of Hanwoo and Yanbian Cattle to identify genome-wide

SNPs and signatures of selection. Mol Cells 38:466–473. https://doi.org/10.14348/molcells.2015.0019
Danilov P, Panchenko D, Tirronen K (2020) Северный олень Восточной Фенноскандии (The reindeer of Eastern Fennoscandia). http://elibrary.krc.karelia.ru/id/eprint/783
Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797. https://doi.org/10.1093/nar/gkh340
Ewels P, Magnusson M, Lundin S, Käller M (2016) MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32:3047–3048. https://doi.org/10.1093/bioinformatics/btw354
Felsenstein J (2005) PHYLIP (phylogeny inference Package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle
Guindon S, Dufayard J-F, Lefort V et al (2010) New algorithms and methods to Estimate Maximum-Likelihood Phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59:307–321. https://doi.org/10.1093/sysbio/syq010
Härkönen S, Bisi J (2007) Suomen metsäpeurakannan hoitosuunnitelma
Heino MT, Salmi A-K, Äikäs T et al (2021) Reindeer from Sámi offering sites document the replacement of wild reindeer genetic lineages by domestic ones in Northern Finland starting from 1400 to 1600 AD. J Archaeol Sci Rep 35:102691. https://doi.org/10.1016/j.jasrep.2020.102691
Iacolina L, Corlatti L, Buzan E et al (2019) Hybridisation in European ungulates: an overview of the current status, causes, and consequences. Mammal Rev 49:45–59. https://doi.org/10.1111/mam.12140
Klokov K (2007) Reindeer husbandry in Russia. Int J Entrep Small Bus 4:726–784. https://doi.org/10.1504/IJESB.2007.014981
Korkeasaari Zoo (2022) Helsinkiin eksynyt orpo vasa varmistui metsäpeuraksi. https://www.korkeasaari.fi/helsinkiin-eksynyt-orpo-vasa-varmistui-metsapeuraksi/ [In Finnish]
Lachance J, Vernot B, Elbers CC et al (2012) Evolutionary history and adaptation from high-coverage whole-genome sequences of diverse African hunter-gatherers. Cell 150:457–469. https://doi.org/10.1016/j.cell.2012.07.009
Lee T-H, Guo H, Wang X et al (2014) SNPhylo: a pipeline to construct a phylogenetic tree from huge SNP data. BMC Genomics 15:162. https://doi.org/10.1186/1471-2164-15-162
Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25:1754–1760. https://doi.org/10.1093/bioinformatics/btp324
Mallet J (2005) Hybridization as an invasion of the genome. Trends Ecol Evol 20:229–237. https://doi.org/10.1016/j.tree.2005.02.010
Montonen M (1974) Suomen Peura, 1st edn. WSOY
Mykrä-Pohja S, Niemi M (2022) Pitkäjänteistä metsäpeuran suojelua - Metsästäjälehti. https://metsastajalehti.fi/riista/pitkajanteista-metsapeuran-suojelua/. Accessed 15 Jun 2023
Niemi M, Rautiainen M, Kilpeläinen P, Turtinen E (2021) Metsäpeuran rotupuhtaustyö Ja Sen kehittäminen 2017–2019. Metsähallitus, Vantaa
Paasivaara A, Hyvärinen M, Timonen P et al (2021) Metsäpeurakanta kasvussa - Metsästäjälehti. https://metsastajalehti.fi/riista/metsapeurakanta-kasvussa/. Accessed 14 Jun 2023
Paasivaara A, Pusenius J, Ruha L, Timonen P (2022) Hirven lentolaskenta – pohjoisen erikoisuus - Metsästäjälehti. https://metsastajalehti.fi/riista/hirven-lentolaskenta-pohjoisen-erikoisuus/. Accessed 14 Jun 2023
Panchenko DV, Tirronen K, Danilov P et al (2017) Оценка численности и распределение лесного северного оленя (Rangifer Taranus Fennicus Lönnb.) в Карелии (Assessment of the forest reindeer (Rangifer Taranus Fennicus Lönnb.) Number and distribution in the Republic of Karelia). Вестник Охотоведения 14:156–165
Panchenko DV, Paasivaara AKJ, Hyvarinen MSA, Krasovskij YA (2021) The wild forest reindeer, Rangifer tarandus fennicus, in the Metsola Biosphere Reserve, Northwest Russia. Nat Conserv Res Заповедная Наука 6:116–126. https://doi.org/10.24189/ncr.2021.026
Pokharel K, Weldenegodguad M, Dudeck S et al (2023) Whole-genome sequencing provides novel insights into the evolutionary history and genetic adaptation of reindeer populations in northern Eurasia. Sci Rep 13:23019. https://doi.org/10.1038/s41598-023-50253-7
Pöllänen AT, Pakanen V-M, Paasivaara A (2023) Survival and cause-specific mortality in adult females of a northern migratory ungulate. Eur J Wildl Res 69:60. https://doi.org/10.1007/s10344-023-01686-y
Purcell S, Neale B, Todd-Brown K et al (2007) PLINK: a Tool Set for whole-genome Association and Population-based linkage analyses. Am J Hum Genet 81:559–575
Rambaut A (2018) FigTree v1. 4.4 (https://github.com/rambaut/figtree/releases/tag/v1.4.4)
Randi E (2005) Management of wild ungulate populations in Italy: Captive-Breeding, hybridisation and genetic consequences of translocations. Vet Res Commun 29:71–75. https://doi.org/10.1007/s11259-005-0025-1
Randi E (2008) Detecting hybridization between wild species and their domesticated relatives. Molecular Ecology, 17.
Rankama T, Ukkonen P (2001) On the early history of the wild reindeer (Rangifer tarandus L.) in Finland. Boreas 30:131–147. https://doi.org/10.1111/j.1502-3885.2001.tb01218.x
Røed KH, Flagstad O, Nieminen M et al (2008) Genetic analyses reveal independent domestication origins of eurasian reindeer. Proc Biol Sci 275:1849–1855. https://doi.org/10.1098/rspb.2008.0332
Røed KH, Kvie KS, Bårdsen B-J et al (2021) Historical and social–cultural processes as drivers for genetic structure in nordic domestic reindeer. Ecol Evol 11:8910–8922. https://doi.org/10.1002/ece3.7728
Schaefer JA, Bergman CM, Luttich SN (2000) Site fidelity of female caribou at multiple spatial scales. Landsc Ecol 15:731–739. https://doi.org/10.1023/A:1008160408257
Somenzi E, Ajmone-Marsan P, Barbato M (2020) Identification of Ancestry informative marker (AIM) panels to assess hybridisation between feral and domestic Sheep. Animals 10:582. https://doi.org/10.3390/ani10040582
Šprem N, Buzan E, Safner T (2023) How we look: European wild mouflon and feral domestic sheep hybrids. https://doi.org/10.1093/cz/zoad031. Curr Zool zoad031
Stothard P, Choi J-W, Basu U et al (2011) Whole genome resequencing of black Angus and Holstein cattle for SNP and CNV discovery. BMC Genomics 12:559–559. https://doi.org/10.1186/1471-2164-12-559
Stroupe S, Forgacs D, Harris A et al (2022) Genomic evaluation of hybridization in historic and modern North American Bison (Bison bison). Sci Rep 12:6397. https://doi.org/10.1038/s41598-022-09828-z
Van der Auwera GA, Carneiro MO, Hartl C et al (2013) From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Current Protocols in Bioinformatics, 11, 11.10.1-11.10.33. https://doi.org/10.1002/0471250953.bi1110s43
Weldenegodguad M, Pokharel K, Ming Y et al (2020) Genome sequence and comparative analysis of reindeer (Rangifer tarandus) in northern Eurasia. Sci Rep 10:8980. https://doi.org/10.1038/s41598-020-65487-y
YLE (2022). Helsingin kaduilla jolkotellut sarvipää selvisi kriittisestä yöstään eläinsairaalassa - poroksi arveltu eläin voikin olla peura. Yle.https://yle.fi/a/3-12262918
Zheng X, Levine D, Shen J et al (2012) A high-performance computing toolset for relatedness and principal component analysis

of SNP data. Bioinforma Oxf Engl 28:3326–3328. https://doi.org/10.1093/bioinformatics/bts606