



OPEN

## Genome-wide association studies for economically important traits in mink using copy number variation

Pourya Davoudi<sup>1</sup>, Duy Ngoc Do<sup>1</sup>, Stefanie Colombo<sup>1</sup>, Bruce Rathgeber<sup>1</sup>, Mehdi Sargolzaei<sup>2,3</sup>, Graham Plastow<sup>4</sup>, Zhiquan Wang<sup>4</sup>, Guoyu Hu<sup>1</sup>, Shafagh Valipour<sup>1</sup> & Younes Miar<sup>1</sup>✉

Copy number variations (CNVs) are structural variants consisting of duplications and deletions of DNA segments, which are known to play important roles in the genetics of complex traits in livestock species. However, CNV-based genome-wide association studies (GWAS) have remained unexplored in American mink. Therefore, the purpose of the current study was to investigate the association between CNVs and complex traits in American mink. A CNV-based GWAS was performed with the ParseCNV2 software program using deregressed estimated breeding values of 27 traits as pseudophenotypes, categorized into traits of growth and feed efficiency, reproduction, pelt quality, and Aleutian disease tests. The study identified a total of 10,137 CNVs (6968 duplications and 3169 deletions) using the Affymetrix Mink 70K single nucleotide polymorphism (SNP) array in 2986 American mink. The association analyses identified 250 CNV regions (CNVRs) associated with at least one of the studied traits. These CNVRs overlapped with a total of 320 potential candidate genes, and among them, several genes have been known to be related to the traits such as *ARID1B*, *APPL1*, *TOX*, and *GPC5* (growth and feed efficiency traits); *GRM1*, *RNASE10*, *WNT3*, *WNT3A*, and *WNT9B* (reproduction traits); *MYO10*, and *LIMS1* (pelt quality traits); and *IFNGR2*, *APEX1*, *UBE3A*, and *STX11* (Aleutian disease tests). Overall, the results of the study provide potential candidate genes that may regulate economically important traits and therefore may be used as genetic markers in mink genomic breeding programs.

American mink (*Neogale vison*) is one of the most important animals in the global fur industry, yet requires highly efficient breeding programs to deal with challenges faced in mink production systems<sup>1</sup>. Several studies have been carried out to discern the genetics of complex traits affecting the sustainability of mink production, including growth<sup>2,3</sup>, feed efficiency<sup>4,5</sup>, disease resistance<sup>6–9</sup>, pelt quality<sup>10–12</sup>, and reproduction traits<sup>13–15</sup>. Understanding the genetic architecture underlying such traits using genome-wide association studies (GWAS) might contribute to expediting the genetic progress through selection, and therefore enhance the production efficiency of the mink industry.

Copy number variations (CNVs) refer to frequently observed structural variations in the form of deletions or duplications greater than 50 base pairs<sup>16</sup>, which cover more of the genome (total bases) and have a higher mutation rate than single nucleotide polymorphisms (SNPs)<sup>17</sup>. Similar to SNPs, CNVs can be applied to detect associations with traits of economic interest in livestock species, and therefore, are considered complementary sources to explain genetic variation contributing to differences in phenotypes<sup>18</sup>. Over the past decade, multiple studies have been carried out to examine the association between CNVs with several important phenotypes in livestock species, such as reproduction<sup>19,20</sup>, health<sup>21–23</sup>, feed efficiency and growth<sup>24–26</sup>, and performance traits<sup>27–29</sup>.

The availability of a high-quality chromosome-based genome assembly<sup>1</sup> and a genome-wide SNP array for American mink facilitates the identification of genetic variations underlying economically important traits. Recently, Davoudi et al.<sup>4</sup> characterized the CNVs in American mink using whole-genome sequencing data. However, up to now, the CNV-based GWAS with economically important traits has not been reported in mink.

<sup>1</sup>Department of Animal Science and Aquaculture, Dalhousie University, Truro, NS, Canada. <sup>2</sup>Department of Pathobiology, University of Guelph, Guelph, ON, Canada. <sup>3</sup>Select Sires Inc., Plain City, OH, USA. <sup>4</sup>Livestock Gentec, Department of Agricultural, Food and Nutritional Science, University of Alberta, Edmonton, AB, Canada. ✉email: miar@dal.ca

Therefore, this study aimed to identify CNV in a large sample of genotyped mink and perform CNV-based GWAS analyses for Aleutian disease tests, growth and feed efficiency, reproduction and pelt quality traits. In addition, we performed functional annotation of the associated CNV regions (CNVRs) to identify the potential candidate genes for these key traits.

## Materials and methods

All procedures applied for this study were approved by the Dalhousie University Animal Care and Use Committee, and we adhered to the Code of Practice for the Care and Handling of Farmed Mink guidelines<sup>30</sup> throughout all phases of the research.

### Sampling, DNA extraction, and genotyping

The current study involved individuals from two separate farms: the Canadian Center for Fur Animal Research (CCFAR) at Dalhousie University, Faculty of Agriculture in Truro, Nova Scotia, Canada, comprising 1411 individuals, and Millbank Fur Farm (MFF) in Rockwood, Ontario, Canada, consisting of 1562 individuals. More details about the animals utilized in this study can be found in our previous work by Hu et al.<sup>31</sup>. Mink were raised in standard farming settings, receiving diets from the byproducts of human food production. These diets were adjusted to meet the specific needs of the animals in each production phase. Detailed information about the feed ingredients, chemical composition, and energy content of these diets during different periods can be found in our prior publication<sup>32</sup>.

DNA was extracted from tongue tissue of the animals utilizing the DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. Evaluation of DNA quantity and quality was conducted using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies Inc., Wilmington, DE), with consistent 260/280 nm readings falling between 1.7 and 2.0 across all samples. After reaching a final concentration of 500 ng, the samples underwent quality checks for DNA and were subsequently genotyped using the Axiom Affymetrix Mink 70K array (Neogen, Lincoln, Nebraska, United States)<sup>33</sup>.

### CNV detection

Initially, the Axiom™ Analysis Suite (Affymetrix™) was applied to perform quality control of raw intensity files and filter genotypes based on dish QC (DQC) values less than 0.82, and a minimum call rate of 97%, following the 'Best Genotyping Practices' Workflow described in Axiom™ Genotyping Solution Data Analysis Guide<sup>34</sup>. The SNPs in sex chromosomes were excluded and only those that passed the quality control were kept for further analyses. The final dataset contained 47,644 SNPs that were located on autosomal chromosomes.

The CNV detection was performed with PennCNV v.1.0.5 software<sup>35</sup>, using the signal intensity ratios (Log R Ratio, LRR) and allelic frequencies (B Allele Frequency, BAF) obtained from the Axiom™ CNV Summary Tool software (Affymetrix™). Then, the PFB (Population Frequency of B allele) file was compiled based on the BAF of each marker in the whole population, using the PennCNV 'compile\_pfb.pl' function. The GC content around each SNP marker is known to affect the signal strength through the potential interference of genomic waves<sup>36</sup>. Therefore, we first estimated the percentage of GC content of 1-Mb genomic regions surrounding each marker (500 kb on each side) using faToTwoBit and hgGcPercent tools provided by UCSC Genome Browser<sup>37</sup> and the FASTA information of the American mink genome assembly<sup>1</sup>. Next, the GC content file was implemented in PennCNV by 'gcmode' function, which applies a regression model for adjusting the high GC content and recovers samples affected by genomic waves<sup>38</sup>. To achieve high-confidence CNV calls, quality control was applied with the following criteria: standard deviations for LRR < 0.35, BAF drift < 0.01, and waviness factor value between -0.05 and 0.05. We only retained those CNVs longer than 1 kb in length including at least three consecutive SNPs located on autosomal chromosomes<sup>28</sup>. Finally, 2063 high-quality samples were kept for subsequent analyses.

### Deregressed EBV values

The deregressed estimated breeding values (dEBVs) were calculated for 27 economically important traits, including 11 growth and feed efficiency traits, eight reproduction traits, five pelt quality traits, and three Aleutian disease tests. The growth and feed efficiency traits include harvest weight (HW), harvest length (HL), final body weight (FBW), final body length (FBL), daily feed intake (DFI), average daily gain (ADG), feed conversion ratio (FCR), residual feed intake (RFI), residual gain (RG), residual intake and gain (RIG), and Kleiber ratio (KR). The reproduction traits include gestation length (GL), total number of kits born (TB), number of kits alive at birth (LB), number of kits alive at weaning (LW), survival rate at birth (SB), average kit weight per litter at birth (AWB), average kit weight per litter at weaning (AWW), and survival rate at weaning (SW). The pelt quality traits include dried pelt size (DPS), overall quality of dried pelt (DQU), dried pelt nap size (DNAP), live grading overall quality of fur (LQU), and live grading nap size (LNAP). The Aleutian disease tests include counterimmunoelectrophoresis (CIEP), the Aleutian mink disease virus (AMDV) capsid protein-based enzyme-linked immunosorbent assay (ELISA-P), and the AMDV antigen-based enzyme-linked immunosorbent assay (ELISA-G). Breeding values were estimated for all individuals using different animal models for growth and feed efficiency described in detail by<sup>4</sup>, for reproduction traits described in detail by<sup>13</sup>, for pelt quality traits described in detail by<sup>10</sup>, and for Aleutian disease tests described in detail by<sup>7</sup>.

The EBV reliabilities were calculated using the following formula:

$$EBV \text{ reliabilities} = 1 - \frac{\text{prediction error variance}}{\text{additive genetic variance of the trait}}$$

Next, the EBV reliabilities were applied to calculate the dEBVs using the method proposed by<sup>39</sup>. The calculations were performed by the ‘wideDRP’ function in DRP package<sup>40</sup> in the R environment<sup>41</sup>, by setting the estimated heritability and the default value of 0.5 for the *c* parameter, which indicates the proportion of genetic variance not explained by markers. The descriptive statistics of the dEBVs for all traits are summarized in Table 1. We removed the animals with dEBV reliability lower than 0.20. The dEBVs were used as the pseudo-phenotype for the association analyses.

### CNV association analysis

ParseCNV2 software<sup>42</sup>, which integrates PLINK<sup>43</sup> for association analyses, was used to detect the association between CNV and dEBVs of the studied traits. ParseCNV2 software converts the CNV events into probe-based statistics for individual CNVs<sup>42</sup>. Since CNV boundaries differ among individuals, it may be difficult to determine the exact start and end points of CNVs, which makes it challenging to classify different CNVs. Therefore, the CNV association tests were conducted for deletions or duplications separately at the probe level. The following model was applied for association testing:

$$y = Xb + e,$$

where *y* is the vector of dEBVs, *X* is the design matrix relating dEBVs to fixed effect of one CNV at a time, *b* is the fixed effect of CNV, and *e* is the vector of random residual effects. The association test output was used to merge neighboring SNPs in proximity (less than 1 Mb apart) with comparable association significance ( $\pm 1$  log p-value) into CNV regions (CNVRs), which constitute a genomic span of at least two consecutive probes. The local lowest P-value for identified probes was chosen to indicate the significant level of the whole CNVRs. To consider multiple testing correction, a threshold less than  $5 \times 10^{-4}$  was applied to consider a CNVR significantly associated with the phenotypes, as proposed by the ParseCNV2 developers<sup>42,44</sup>.

### Gene annotation

The list of genes in the latest American mink reference genome (ASM\_NN\_V1)<sup>1</sup> was downloaded from the NCBI and the ‘intersect’ function in Bedtools<sup>45</sup> was used to detect the genes that overlapped with significant

Trait	Abbreviations	Numbers	Mean	SD	dEBVs	
					Min	Max
Harvest weight	HW	1985	-0.01	0.30	-1.82	1.92
Harvest length	HL	1921	0.10	2.03	-7.92	9.91
Final body weight	FBW	1037	0.00	0.30	-1.15	1.37
Final body length	FBL	1038	-0.08	1.98	-6.59	5.83
Daily feed intake	DFI	1872	0.01	0.07	-0.15	0.23
Average daily gain	ADG	1044	-0.02	2.19	-10.14	9.18
Feed conversion ratio	FCR	1036	0.34	8.60	-23.39	61.52
Residual feed intake	RFI	1044	0.00	0.02	-0.12	0.14
Residual gain	RG	1042	-0.04	1.51	-7.19	5.70
Residual intake and gain	RIG	1043	-0.04	1.48	-6.76	6.43
Kleiber ratio	KR	1044	-0.01	1.15	-6.05	4.64
Counterimmunoelectrophoresis test	CIEP	1356	0.00	0.12	-0.63	0.34
VP2 based enzyme-linked immunosorbent assay test	ELISA-P	1356	0.04	2.12	-5.16	10.34
AMDV-G based enzyme-linked immunosorbent assay test	ELISA-G	1356	-0.10	2.23	-6.39	10.83
Gestation length	GL	1321	-0.44	2.74	-29.01	25.43
Total number of kits born	TB	1321	0.25	1.37	-6.51	11.10
Number of kits alive at birth	LB	1319	0.35	1.26	-5.15	9.14
Number of kits alive at weaning	LW	1314	-0.04	1.20	-8.51	7.39
Survival rate at birth	SB	1169	2.39	7.77	-61.69	46.51
Average kit weight per litter at birth	AWB	1168	0.01	0.83	-4.90	5.98
Average kit weight per litter at weaning	AWW	1167	0.95	23.02	-168.37	132.25
Survival rate at weaning	SW	1166	-0.27	14.26	-86.29	51.27
Dried pelt size	DPS	1169	0.02	0.70	-7.20	8.24
Overall quality of dried pelt	DQU	1148	-0.01	0.56	-3.92	2.17
Dried pelt nap size	DNAP	1159	0.17	1.10	-4.91	4.94
Live grading overall quality of fur	LQU	1260	0.05	0.56	-3.96	3.79
Live grading nap size	LNAP	1260	0.14	0.79	-3.84	3.60

**Table 1.** Descriptive statistics of the deregressed EBV (dEBVs) for growth, feed efficiency, Aleutian disease tests, pelt quality and reproduction traits in American mink.

CNV regions. Finally, an extensive review of the literature was performed to investigate the biological function of identified candidate genes.

## Results

### CNV identification and distribution

Using the PennCNV software based on the Hidden Markov Model method<sup>35</sup>, a total of 10,137 CNV events were identified from 2063 individuals that passed the quality control criteria (Supplementary Table S1). While PennCNV is extensively utilized for CNV detection in genotyping array data, it is essential to note its limitations. The internal HMM model applied in the software specifically considers successive SNPs at each step, making it particularly sensitive to local noise. This sensitivity often results in false positives, over-segmentation (where a true CNV is incorrectly divided into smaller segments), and generally imprecise boundaries in the PennCNV calls<sup>46</sup>. Among the total identified CNVs, 6968 (68.74%) were duplications and 3169 (31.26%) were deletions, with a deletions/duplications CNV ratio of 0.45. The length of the CNV events ranged from 1.05 to 6148.34 kb, with an average size of 109.69 kb. Table 2 presents the descriptive statistics of the identified CNVs in the American mink genome. Analysis of the distribution of CNV size showed that approximately half of the CNVs ranged from 1 to 50 kb, with relatively rare CNV events (3.97%) larger than 500 kb (Fig. 1a). The number of CNVs on each chromosome and the chromosome length showed a strong positive linear correlation (Fig. 1c,  $r=0.78$ ), such that 1282 CNVs were identified for the largest chromosome (Chromosome 1) and 225 CNVs for the smallest chromosome (chromosome 14; Fig. 1b).

### Association analyses

In order to explore the effect of CNVs on the complex phenotypes, CNV-based association analyses were carried out for the 27 economically important traits in American mink. Association analyses revealed that 250 CNVRs (71 deletions and 179 duplications) were significantly associated with at least one of the studied traits ( $P < 0.0005$ ). Manhattan plots for significant CNVRs across the autosomes associated with all studied traits are shown in Figs. 2, 3, 4, 5. These significant regions were identified across all 14 autosomes, while chromosome one showed the largest number ( $n = 53$ ).

The overview of the top significant CNVRs associated with each studied trait is shown in Table 3. Detailed information for all significant CNVRs with their overlapped candidate genes is provided in Supplementary Table S2. The highest number of significant CNVRs ( $n = 27$ ) were associated with TB, comprising the most significant region (ID: CNVR54) with a p-value of  $3.58 \times 10^{-14}$ . In addition, the average length of significant CNVRs was 66.2 kb, ranging from 1.23 to 444.54 kb.

### Candidate genes within the significant CNVR

We further investigated the candidate genes encompassing the significant CNVRs. The results revealed that a total of 320 potential candidate genes overlapped with significant CNVRs based on the annotation of the American mink genome (Supplementary Table S2). The duplication CNVR on chromosome 7 (ID: CNVR143) overlapped with the highest number of genes ( $n = 13$ ) while no genes identified within 80 significant CNVRs.

Using the information from the GeneCards database and an extensive literature review, several candidate genes were found to be related to growth and feed efficiency traits (*ARID1B*, *APPL1*, *TOX*, and *GPC5*), reproduction traits (*GRM1*, *RNASE10*, *WNT3*, *WNT3A*, and *WNT9B*), pelt quality traits (*MYO10*, and *LIMS1*), and Aleutian disease tests (*IFNGR2*, *APEX1*, *UBE3A*, and *STX11*).

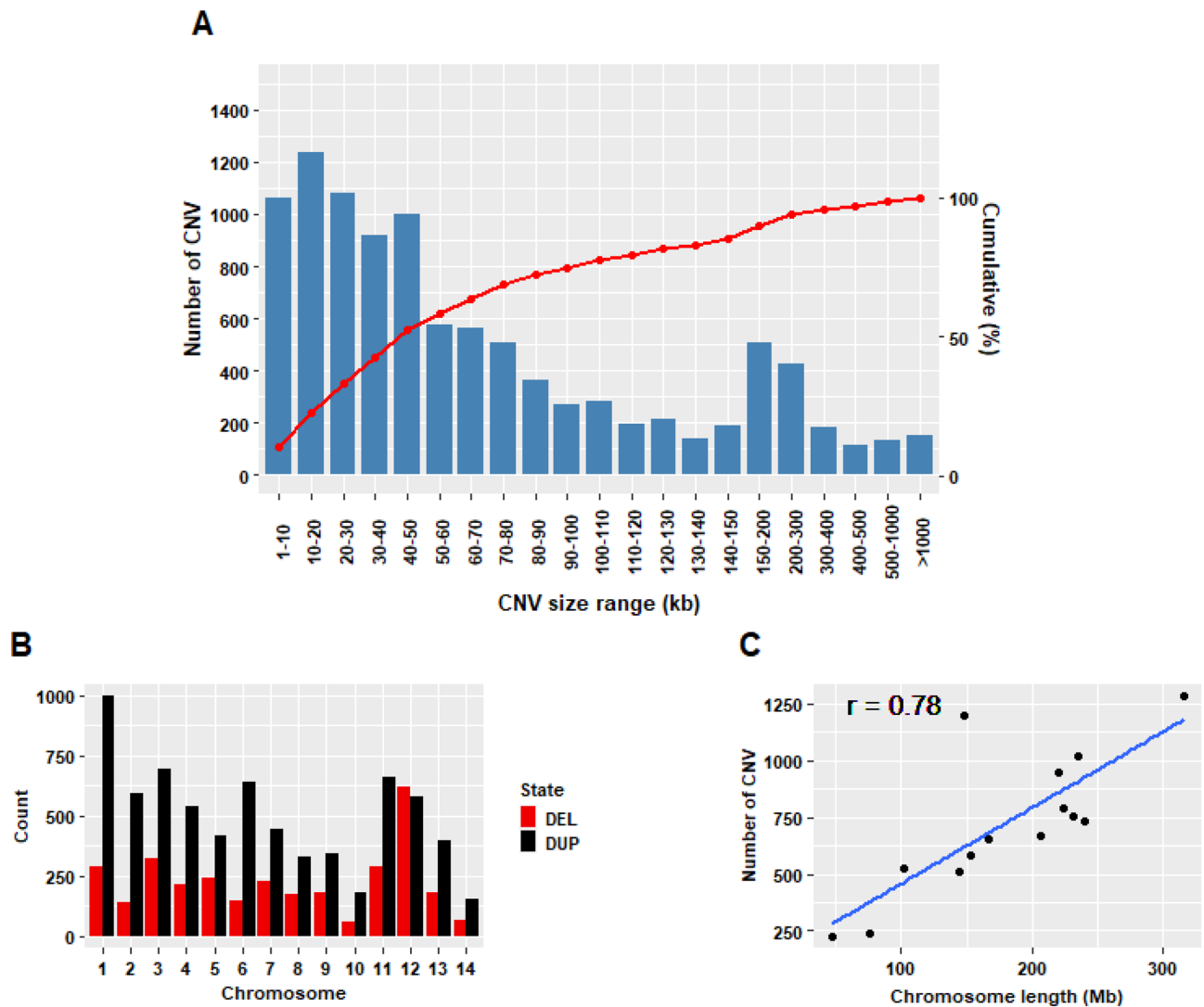
## Discussion

Genome-wide association studies using SNP markers have been instrumental in unraveling the underpinning of complex traits<sup>47</sup>. In recent years, CNVs have gained widespread utilization as a supplementary tool in association studies, adding in the identification of genetic variants associated with economically important traits and shedding light on the elucidating the genetic basis of these traits across different livestock species<sup>28,48–51</sup>. To the best of our knowledge, there is no prior research had delved into the realm of CNV associations with diverse phenotypes in American mink.

We conducted the CNV-based association studies using Affymetrix Mink 70K SNP array to identify potential genetic variants associated with dEBVs of 27 different traits such as growth and feed efficiency, reproduction, Aleutian disease tests, and pelt quality traits. In total, 10,137 CNVs were identified, with an average number of five CNVs per sample. Although the average number of detected CNVs per individual is substantially less than our previous study using whole-genome sequencing data (average number of 1647.3), it is in agreement with the results of other studies that used SNP genotyping data with a similar marker density<sup>21,52–54</sup>. It is well-known

CNV	Number	Mean	Length (bp)	
			Minimum	Maximum
Deletion	3169	98,984.12	1049	2,956,427
Duplication	6968	114,555.32	1128	6,148,335
Overall	10,137	109,687.54	1049	6,148,335

**Table 2.** Descriptive statistics of CNVs detected in American mink genome.

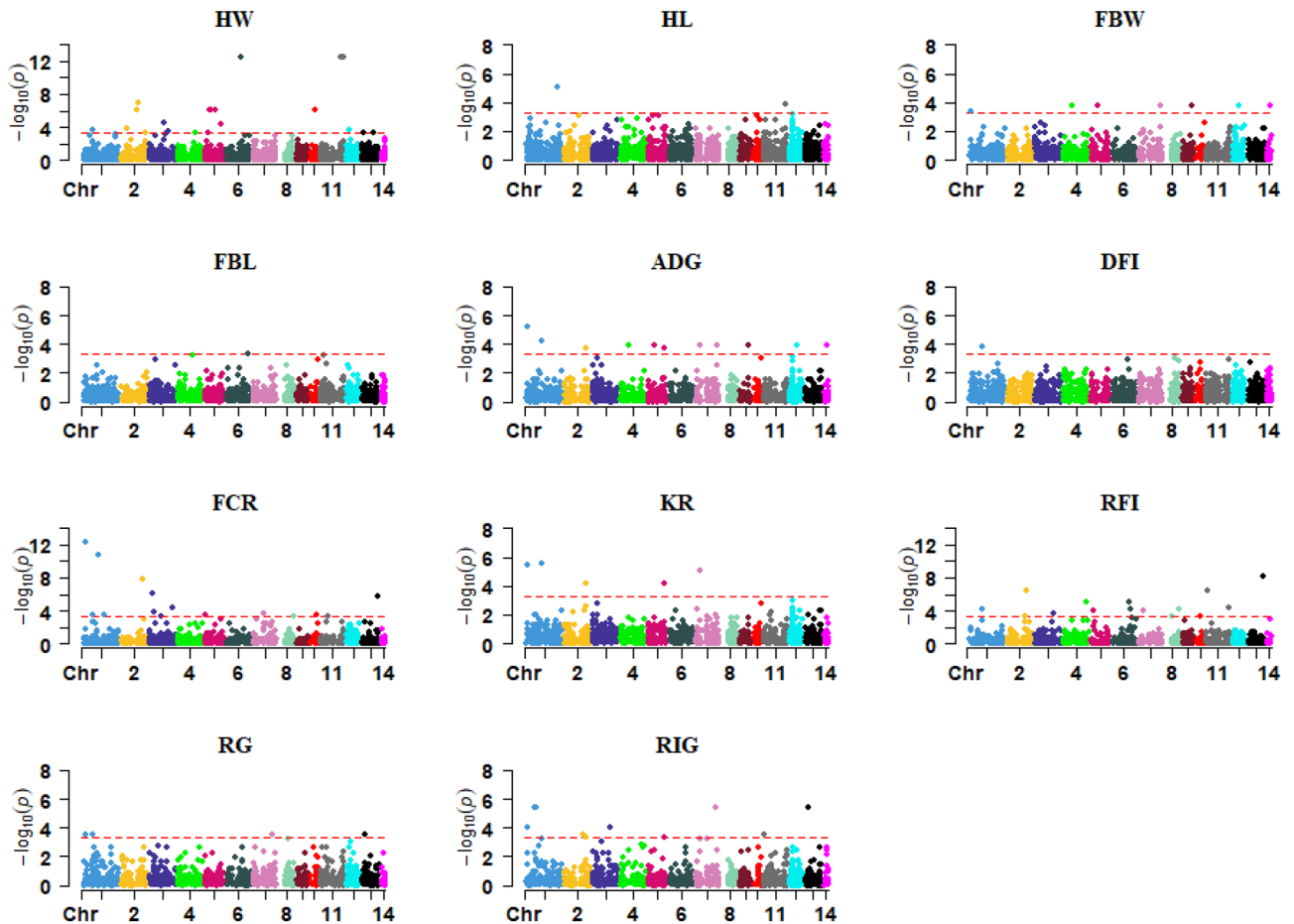


**Figure 1.** Graphical representation of identified CNVs: (A) Distribution of CNV sizes, (B) numbers of CNVs identified across autosomal chromosomes, (C) correlation between CNV numbers and chromosome length.

that the SNP genotyping density affects the number and length of the identified CNVs<sup>21</sup>. The average length of identified CNVs (109.69 kb) is much longer than our previous study with an average size of 7.4 kb, showing differences in resolution and coverage of genome between SNP array and whole genome sequencing data, yet falls within the range of other studies using comparable SNP array datasets<sup>55,56</sup>. A total of 250 significant CNVRs were associated with at least one of the studied traits ( $P < 0.0005$ ), overlapping with 320 potential candidate genes.

For growth and feed efficiency traits, we identified 86 CNVRs associated with eleven traits. Within these significant CNV segments, we identified *ARID1B*, *APPL1*, *TOX*, and *GPC5* genes, which might have large impacts on growth rate and feed efficiency in American mink. The *ARID1B* gene is overlapped with the duplication CNVR1 (Chr1:10,430,780–10,527,641), which was significantly associated with traits such as FBW, ADG, FCR, KR, and RIG. The *ARID1B* gene, which plays a key role in controlling the maturation of neurons during brain development<sup>57</sup>, is the commonly mutated gene in Coffin-Siris syndrome, a genetic disorder characterized by intellectual disability, developmental delay, and growth impairment<sup>58,59</sup>. Yu et al.<sup>60</sup> reported that the *ARID1B* gene overlapped with identified CNVs in patients with short stature and developmental disorder, indicating the critical function of *ARID1B* mutations in human height regulation. Interestingly, Bovo et al.<sup>61</sup> detected a region being targeted by selection pressure, harboring the *ARID1B* gene in different pig breeds that grouped by their size, supporting the effect of this gene on body size. The *APPL1* gene located on duplication CNVR125 (Chr6:198,790,330–198,804,380) interacts with several proteins such as adiponectin receptors, AMPK, and Rab5 (a small GTPase downstream of *APPL1*) to regulate apoptosis, cell proliferation, metabolism and insulin sensitivity in energy homeostasis, resulting in increased glucose uptake and fatty acid oxidation<sup>62</sup>. Schweer et al.<sup>63</sup> reported that the *APPL1* gene is associated with feed efficiency traits in beef cattle through the regulation of glucose.

The *TOX* gene, located in the duplication CNVR75 (Chr4:75,332,844,330–75,473,471), which overlapped with the identified CNVR in our previous study in American mink<sup>64</sup>, was significantly associated with FBW and ADG. The *TOX* gene is a family member of high-mobility group box proteins and serves as a regulator of

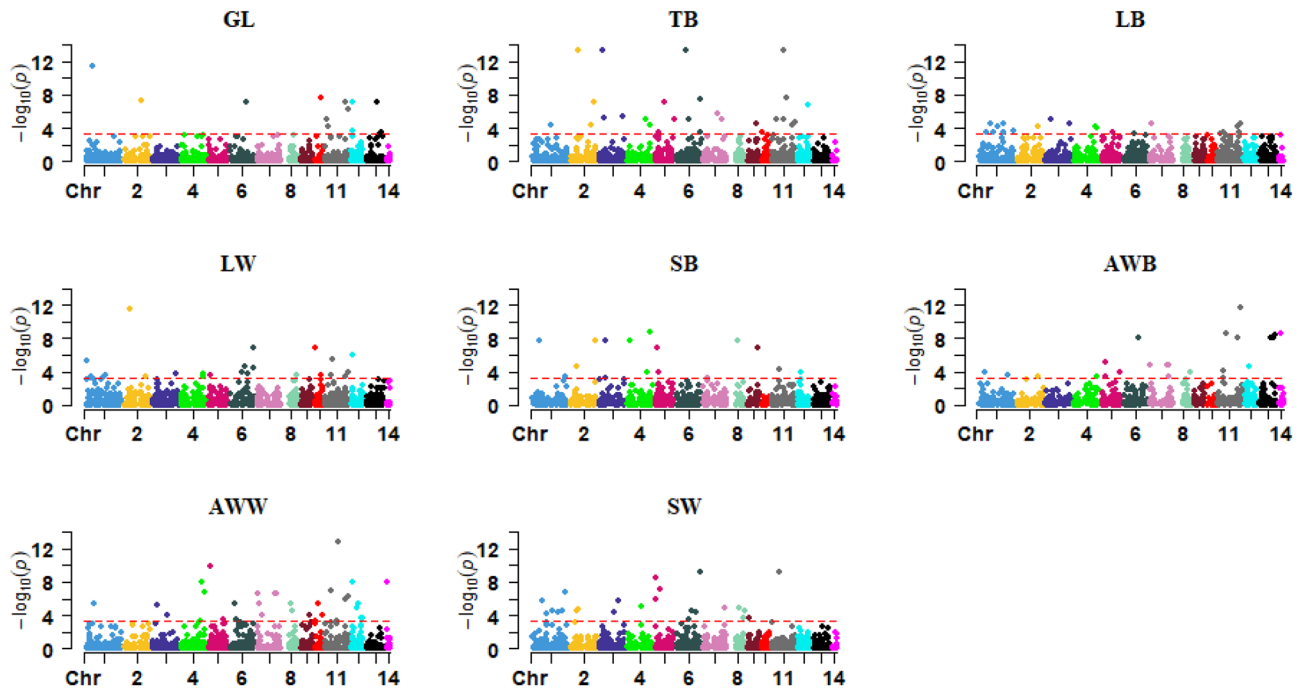


**Figure 2.** Manhattan plots for CNV regions across the 14 autosomal chromosomes associated with feed efficiency and growth traits. The horizontal line in each plot represents the threshold for significance ( $P < 0.0005$ ) suggested by ParseCNV2 developers. HW: harvest weight; HL: harvest length; FBW: final body weight; FBL: final body length; ADG: average daily gain; DFI: daily feed intake; FCR: feed conversion ratio; KR: Kleiber ratio; RFI: residual feed intake; RG: residual gain; RIG: residual intake and gain.

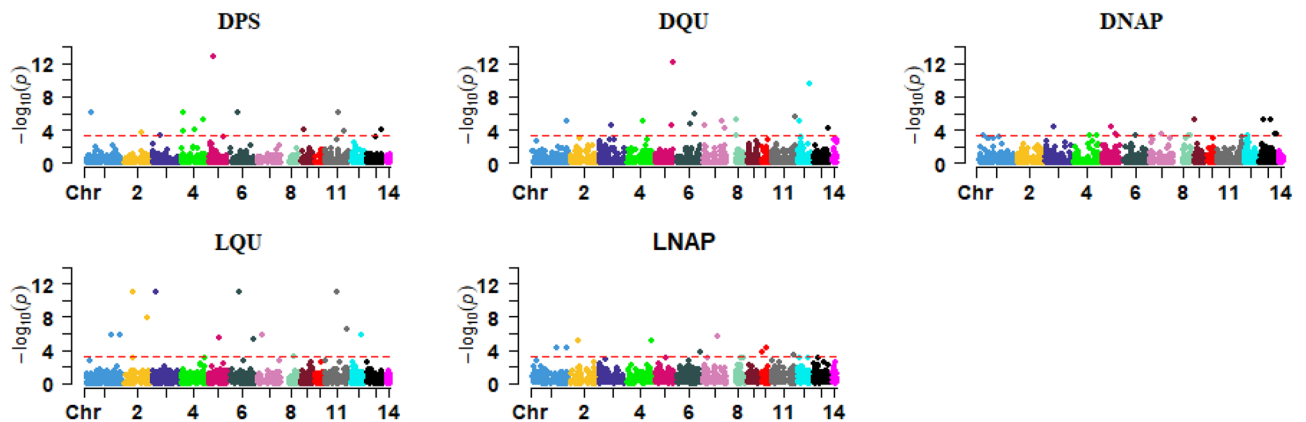
gene expression, mostly through modifying the density of the chromatin structure<sup>65</sup>. In cattle, numerous studies demonstrated that the *TOX* gene is associated with feed efficiency<sup>66</sup>, growth<sup>67</sup>, carcass traits<sup>68–70</sup>, and development of puberty<sup>71</sup>. Furthermore, it is shown that the *TOX* gene is associated with weight gain, obesity, and metabolic syndrome-related phenotypes in humans<sup>72</sup>. The *GPC5* gene, which mediates several functions in the control of cell division and growth regulation<sup>73</sup>, is found within a deletion CNVR96 (Chr5:149,166,233–149,199,404), associating with ADG, RIG, and KR. Congruent with these findings, *GPC5* was reported as a candidate gene located within a significant SNP with effects on RIG, RFI, and efficiency of intake in beef cattle<sup>74</sup>. Moreover, other GWAS indicated the association of the *GPC5* gene with human height<sup>75</sup>, body mass index<sup>76</sup>, and body size/body weight in chicken<sup>77,78</sup>.

Regarding the female reproduction traits, we found 168 significant CNVRs, overlapping with several functional genes, among which *GRM1*, *RNASE10*, *WNT3*, *WNT3A*, and *WNT9B* might be the candidate genes related to female reproduction in mink. The *GRM1* gene, which was previously identified in a CNV study in American mink<sup>64</sup>, was located within the duplication CNVR15 (Chr1: 56,870,177–56,925,961), associated with AWB trait. The *GRM1* was a gene of interest reported in several studies to be associated with female reproduction in different livestock species, such as seasonal reproduction in sheep<sup>79</sup>, litter size in goats<sup>80</sup>, number of teats and litter traits in pig<sup>81</sup>, and fertility-related traits in cattle<sup>82</sup>. Interestingly, it was reported that the *GRM1* gene located within structural variations and runs of homozygosity regions associated with litter traits in pigs<sup>83,84</sup>, highlighting the hypothesis that this gene might be a candidate gene for female reproduction in American mink. It was suggested that *RNASE10* gene action in the proximal epididymis is vital for the acquisition of spermatozoa adhesiveness, eventually affecting the mode of sperm transport in the female reproductive tract<sup>85</sup>.

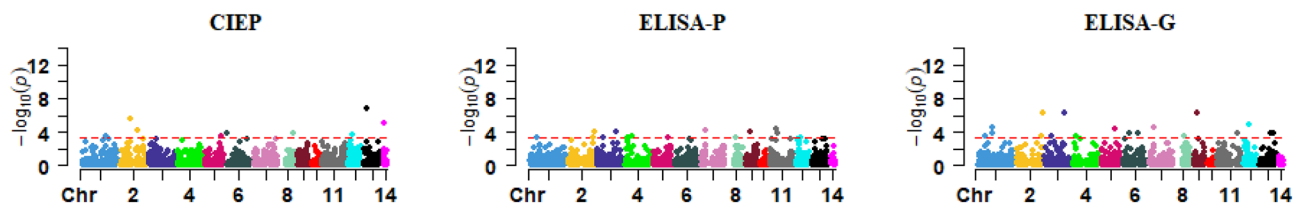
In the current study, several *WNT* family genes were identified to be associated with reproduction traits in American mink, including *WNT3* and *WNT9B* (both found within the duplication CNVR118), and *WNT3A* (overlapped with the deletion CNVR23). It is well-documented that the expression of *WNT3* during the early pregnancy mediates the stromal cell proliferation and trophoblast invasion, eventually affecting the embryonic development<sup>86</sup>. Comparably, the *WNT9B* gene has been reported as one of the key genes associated with inducing the gonadotropin-releasing hormone secretion during follicular development in sheep<sup>87</sup>. Another gene, *WNT3A*,



**Figure 3.** Manhattan plots for CNV regions across the 14 autosomal chromosomes associated with reproduction traits. The horizontal line in each plot represents the threshold for significance ( $P < 0.0005$ ) suggested by ParseCNV2 developers. GL: gestation length; TB: total number of kits born; LB: number of kits alive at birth; LW: number of kits alive at weaning; SB: survival rate at birth; AWB: average kit weight per litter at birth; AWW: average kit weight per litter at weaning; SW: survival rate at weaning.



**Figure 4.** Manhattan plots for CNV regions across the 14 autosomal chromosomes associated with pelt quality traits. The horizontal line in each plot represents the threshold for significance ( $P < 0.0005$ ) suggested by ParseCNV2 developers. DPS: dried pelt size; DQU: overall quality of dried pelt; DNAP: dried pelt nap size; LQU: live grading overall quality of fur; LNAP: live grading nap size.



**Figure 5.** Manhattan plots for CNV regions across the 14 autosomal chromosomes associated with Aleutian disease tests. The horizontal line in each plot represents the threshold for significance ( $P < 0.0005$ ) suggested by ParseCNV2 developers. CIEP: counterimmunoelectrophoresis test; ELISA-P: VP2 based enzyme-linked immunosorbent assay test; ELISA-G: AMDV-G based enzyme-linked immunosorbent assay test.

CNVR ID	Type	Chromosome	Start position	End position	Length (bp)	Associated traits	P-value	Candidate genes
CNVR54	Duplication	2	76,047,394	76,080,498	33,104	TB	3.58E-14	–
CNVR112	Deletion	5	45,720,950	45,756,856	35,906	DPS	1.26E-13	<i>ITGA3</i> , <i>LOC122907723</i>
CNVR203	Duplication	11	119,613,197	119,626,953	13,756	AWW	1.44E-13	<i>LOC122890709</i> , <i>LOC122889431</i>
CNVR126	Duplication	6	143,324,461	143,377,043	52,582	HW	3.48E-13	–
CNVR1	Duplication	1	10,430,780	10,527,641	96,861	FCR	5.35E-13	<i>ARID1B</i>
CNVR113	Deletion	5	158,304,905	158,351,587	46,682	DQU	6.26E-13	<i>METTL21C</i> , <i>BIVM</i> , <i>CCDC168</i>
CNVR200	Duplication	11	208,363,505	208,379,126	15,621	AWB	2.24E-12	<i>DLGAP2</i>
CNVR56	Duplication	2	51,828,542	51,834,524	5982	LW	2.95E-12	–
CNVR20	Duplication	1	48,882,466	48,894,690	12,224	GL	3.08E-12	<i>BCKDHB</i>
CNVR54	Duplication	2	76,047,394	76,080,498	33,104	LQU	8.35E-12	–
CNVR216	Deletion	11	71,374,870	71,389,512	14,642	SW	5.42E-10	<i>TBC1D9</i>
CNVR90	Duplication	4	201,592,245	201,635,594	43,349	SB	1.71E-09	<i>KCND2</i>
CNVR238	Duplication	13	139,238,384	139,379,378	140,994	RFI	7.32E-09	–
CNVR235	Duplication	13	26,765,710	26,829,160	63,450	CIEP	1.98E-07	–
CNVR40	Duplication	2	237,951,459	238,012,681	61,222	ELISA_G	5.81E-07	–
CNVR159	Deletion	7	130,788,259	130,823,564	35,305	LNAP	1.9E-06	<i>PIWIL4</i>
CNVR2	Duplication	1	133,977,608	134,011,851	34,243	KR	2.64E-06	<i>SYCP2L</i>
CNVR13	Duplication	1	75,139,496	75,250,132	110,636	RIG	3.97E-06	–
CNVR1	Duplication	1	10,430,780	10,527,641	96,861	ADG	6.29E-06	<i>ARID1B</i>
CNVR179	Duplication	9	12,654,424	12,657,997	3573	DNAP	6.66E-06	<i>TLL11</i>
CNVR70	Duplication	3	46,223,063	46,323,843	100,780	LB	7.80E-06	<i>NDUFA10</i>
CNVR9	Duplication	1	282,031,840	282,086,441	54,601	HL	9.33E-06	–
CNVR192	Duplication	11	56,914,479	56,922,804	8325	ELISA_P	4.19E-05	<i>LOC122890225</i>
CNVR4	Duplication	1	119,299,899	119,337,598	37,699	DFI	1.53E-04	<i>HSD17B8</i> , <i>SLC39A7</i> , <i>RXRβ</i> , <i>COL11A2</i>
CNVR75	Duplication	4	75,332,844	75,473,471	140,627	FBW	1.81E-04	<i>TOX</i>
CNVR12	Duplication	1	11,780,319	11,833,194	52,875	RG	2.84E-04	<i>NOX3</i>
CNVR125	Duplication	6	198,790,330	198,804,380	14,050	FBL	4.27E-04	<i>ASB14</i> , <i>DNAH12</i> , <i>APPL1</i>

**Table 3.** Overview of the top significant CNVRs associated with all studied traits in American mink. TB: total number of kits born, DPS: dried pelt size, AWW: average kit weight per litter at weaning, HW: harvest weight, FCR: feed conversion ratio, DQU: overall quality of dried pelt, AWB: average kit weight per litter at birth, LW: number of kits alive at weaning, GL: gestation length, LQU: live grading overall quality of fur, SW: survival rate at weaning, SB: survival rate at birth, RFI: residual feed intake, CIEP: counterimmunoelectrophoresis test, ELISA\_G: AMDV-G based enzyme-linked immunosorbent assay test, LNAP: Live grading nap size, KR: Kleiber ratio, RIG: residual intake and gain, ADG: average daily gain, DNAP: dried pelt nap size, LB: number of kits alive at birth, HL: harvest length, ELISA\_P: VP2 based enzyme-linked immunosorbent assay test, DFI: daily feed intake, FBW: final body weight, RG: residual gain, FBL: final body length.

is known as the main regulator of reproductive behavior and follicular activity associated with estrus, which in turn may contribute to the reproductive efficiency in cattle<sup>88</sup>.

The gene annotation within significant CNVRs for pelt quality traits identified some functional candidate genes affecting fur characteristics, such as *MYO10* and *LIMS1*. The *MYO10* gene is an integral member of the myosin family, which is involved in various cellular processes such as dynamic actin remodeling, cell migration and adhesion, and filopodia formation<sup>89,90</sup>. Our findings are in accordance with previous studies indicating that the *MYO10* gene plays a key role in mediating skin pigmentation through regulating melanosome transportation in the skin<sup>91–93</sup>. It was demonstrated that the melanocytes present in the skin control the quantity and types of melanosomes, ultimately determining the coat color<sup>94</sup>. Notably, it was shown that the *MYO10* gene mutation altered the coat color pigmentation pattern in mice, further supporting its role in facilitating melanoblast migration<sup>95</sup>. Interestingly, another member of the myosin superfamily, the *MYO5A* gene, has been widely documented for its impact on coat color phenotype in different species<sup>96–98</sup>, and specifically in American mink<sup>64,99</sup>.

The *LIMS1* gene is located within the duplication CNVR170 (Chr8: 65,501,492–65,735,953), associated with the DQU trait. The *LIMS1* gene involved in the control of cell signaling, adhesion, migration, proliferation, and survival<sup>100</sup>. Several studies demonstrated that the *LIMS1* gene regulates cell adhesion and spreading through



the ternary protein complex of integrin-linked kinase (ILK), PINCH, and parvin<sup>101</sup>. To this end, ILK has been reported as a crucial factor for hair morphogenesis<sup>102</sup>. Interestingly, Endo et al.<sup>103</sup> reported that the *LIMS1* gene was associated with hair morphology and density in East Asians. Furthermore, it was shown that the loss of *LIMS1* gene expression from mouse keratinocytes resulted in impaired hair follicle growth<sup>104</sup>, supporting the importance of this gene on fur development in mink.

Aleutian mink disease virus causes autoimmune disorders in mink by stimulating their immune response to produce antibodies and form immune complexes<sup>8,105</sup>. For CIEP, ELISA-P, and ELISA-G, we identified 9, 12, and 22 significant CNVRs, respectively, which overlapped with several immune-related genes such as *IFNGR2*, *APEX1*, *UBE3A*, and *STX11*. The *IFNGR2* gene encodes IFN $\gamma$ R2, which is part of the IFN- $\gamma$  receptor complex that is overexpressed in an inflammatory environment<sup>106</sup>. It is well established that *IFNGR2* is an important regulator for IFN- $\gamma$ -STAT1 signaling in T cells<sup>107</sup>, in turn, the dysregulation of the *IFNGR2* gene is associated with a variety of autoimmune diseases<sup>108</sup>. The *APEX1* and *UBE3A* genes overlapped with the duplication CNVR236 (Chr13: 96,258,171–96,346,188) and the deletion CNVR237 (Chr13: 125,264,701–125,304,748), respectively, both associated with ELISA\_G, GL and AWB traits. In agreement with our results, Hu et al.<sup>7</sup> indicated a favorable genetic correlation between Aleutian disease test and reproduction traits in American mink, which suggested the potential for genetic selection of Aleutian disease test traits to alleviate the adverse impact caused by Aleutian disease in mink farms.

The *APEX1* gene (also called *APE1*) encodes a multifunctional protein that regulates the DNA base excision repair and redox activities, the latter demonstrated to be involved in mediating the T helper cell 1 (Th1) response<sup>109</sup>. In addition, it is well-documented that the *APEX1* gene plays a proinflammatory function in stimulating cytokine and chemokine expression, eventually contributing to innate and adaptive immunity processes<sup>110</sup>. It has been confirmed that the *UBE3A* gene, present in both glutamatergic and GABAergic neurons in the brain, functions as a transcriptional regulator of the immune system within the brain<sup>111</sup>. Recently, Zhang et al.<sup>112</sup> revealed that the *UBE3A* gene within a deletion CNV is associated with the enrichment levels of immune signaling pathways, eventually enhancing antitumor immunity and immunogenicity. The *STX11* gene, which is a member of the SNARE family, is highly expressed in immune tissues such as the thymus, spleen, and lymph nodes, regulating the IFN- $\gamma$  secretion from natural killer cells, consequently mediating the immune cell function<sup>113,114</sup>.

## Conclusion

For the first time in American mink, the CNV-based GWAS were applied for economically important traits using the Affymetrix Mink 70K SNP array. We identified 10,137 CNVs, including 6968 duplications and 3169 deletions, among which 250 CNVRs were significantly associated with at least one trait. From this, we identified several candidate genes contributing to the growth and feed efficiency (*ARID1B*, *APPL1*, *TOX*, and *GPC5*), reproduction (*GRM1*, *RNASE10*, *WNT3*, *WNT3A*, and *WNT9B*), pelt quality (*MYO10*, and *LIMS1*), and Aleutian disease tests (*IFNGR2*, *APEX1*, *UBE3A*, and *STX11*). Overall, the associated CNVRs and respective candidate genes in the current study supply additional information, complementary to GWAS analyses solely based on SNP markers, further helping reveal the genetic basis of traits of economic interest in American mink.

## Data availability

The datasets used and analyzed during the current study may be available from the corresponding author on academic request.

Received: 4 October 2023; Accepted: 20 December 2023

Published online: 02 January 2024

## References

- Karimi, K. et al. A chromosome-level genome assembly reveals genomic characteristics of the American mink (*Neogale vison*). *Commun. Biol.* **5**, 1381 (2022).
- Do, D. N., Hu, G., Salek Ardestani, S. & Miar, Y. Genetic and phenotypic parameters for body weights, harvest length, and growth curve parameters in American mink. *J. Anim. Sci.* **99**, 1–7 (2021).
- Madsen, M. et al. Combined analysis of group recorded feed intake and individually recorded body weight and litter size in mink. *Animal* **14**, 1793–1801 (2020).
- Davoudi, P. et al. Genetic and phenotypic parameters for feed efficiency and component traits in American mink. *J. Anim. Sci.* **100**, 1–10 (2022).
- Shirali, M., Nielsen, V., Møller, S. & Jensen, J. Longitudinal analysis of residual feed intake and BW in mink using random regression with heterogeneous residual variance. *Animal* **9**, 1597–1604 (2015).
- Hu, G. et al. Genetic and phenotypic correlations between Aleutian disease tests with body weight, growth, and feed efficiency traits in mink. *J. Anim. Sci.* **100**, 1–14 (2022).
- Hu, G., Do, D. N., Karimi, K. & Miar, Y. Genetic and phenotypic parameters for Aleutian disease tests and their correlations with pelt quality, reproductive performance, packed-cell volume, and harvest length in mink. *J. Anim. Sci.* **99**, 1–12 (2021).
- Karimi, K., Farid, A. H., Myles, S. & Miar, Y. Detection of selection signatures for response to Aleutian mink disease virus infection in American mink. *Sci. Rep.* **11**, 2944 (2021).
- Farid, A., Daftarian, P. & Fatehi, J. Transmission dynamics of Aleutian mink disease virus on a farm under test and removal scheme. *J. Vet. Sci. Med. Diagn.* **7**, 2–10 (2018).
- Valipour, S. et al. Genetic and phenotypic parameters for pelt quality and body length and weight traits in American mink. *Animals* **12**, 3184 (2022).
- Valipour, S. et al. Genome-wide detection of selection signatures for pelt quality traits and coat color using whole-genome sequencing data in American mink. *Genes* **13**, 1939 (2022).
- Thirstrup, J., Jensen, J. & Lund, M. Genetic parameters for fur quality graded on live animals and dried pelts of American mink (*Neovison vison*). *J. Anim. Breed. Genet.* **134**, 322–331 (2017).
- Karimi, K., Sargolzaei, M., Plastow, G. S., Wang, Z. & Miar, Y. Genetic and phenotypic parameters for litter size, survival rate, gestation length, and litter weight traits in American mink. *J. Anim. Sci.* **96**, 2596–2606 (2018).

14. Kołodziejczyk, D. & Socha, S. Analysis of effectiveness of breeding work and estimation of genetic and phenotypic trends for reproductive traits in American mink. *Ann. Anim. Sci.* **11**, 273–282 (2011).
15. Hansen, B., Su, G. & Berg, P. Genetic variation in litter size and kit survival of mink (*Neovison vison*). *J. Anim. Breed. Genet.* **127**, 442–451 (2010).
16. Mills, R. E. *et al.* Mapping copy number variation by population-scale genome sequencing. *Nature* **470**, 59–65 (2011).
17. Geistlinger, L. *et al.* Widespread modulation of gene expression by copy number variation in skeletal muscle. *Sci. Rep.* **8**, 1399 (2018).
18. Hay, E. H. A. *et al.* Genomic predictions combining SNP markers and copy number variations in Nellore cattle. *BMC Genomics* **19**, 1–8 (2018).
19. Liu, M. *et al.* Array CGH-based detection of CNV regions and their potential association with reproduction and other economic traits in Holsteins. *BMC Genomics* **20**, 1–10 (2019).
20. Zheng, X. *et al.* CNV analysis of Meishan pig by next-generation sequencing and effects of AHR gene CNV on pig reproductive traits. *J. Anim. Sci. Biotechnol.* **11**, 1–11 (2020).
21. Butty, A. M. *et al.* Genome-wide association study between copy number variants and hoof health traits in Holstein dairy cattle. *J. Dairy Sci.* **104**, 8050–8061 (2021).
22. Berton, M. P. *et al.* Genomic integration to identify molecular biomarkers associated with indicator traits of gastrointestinal nematode resistance in sheep. *J. Anim. Breed. Genet.* **139**, 502–516 (2022).
23. Schurink, A. *et al.* Copy number variations in Friesian horses and genetic risk factors for insect bite hypersensitivity. *BMC Genet.* **19**, 1–13 (2018).
24. Li, M., Yin, C., Zhao, F. & Liu, Y. Copy number variation association studies for sheep tail-relevant traits in Hulunbair sheep. *Anim. Genet.* **53**, 897–900 (2022).
25. Wang, Z. *et al.* Genome-wide detection of CNVs and association with body weight in sheep based on 600K SNP arrays. *Front. Genet.* **11**, 558 (2020).
26. Zhou, Y. *et al.* Genome-wide copy number variant analysis reveals variants associated with 10 diverse production traits in Holstein cattle. *BMC Genomics* **19**, 1–9 (2018).
27. Yang, L. *et al.* Genomic sequencing analysis reveals copy number variations and their associations with economically important traits in beef cattle. *Genomics* **113**, 812–820 (2021).
28. Fernandes, A. C. *et al.* Genome-wide detection of CNVs and their association with performance traits in broilers. *BMC Genomics* **22**, 1–18 (2021).
29. Ding, R. *et al.* A composite strategy of genome-wide association study and copy number variation analysis for carcass traits in a Duroc pig population. *BMC Genomics* **23**, 1–16 (2022).
30. NFACC. (National Farm Animal Care Council, 2013).
31. Hu, G. *et al.* Population genomics of American mink using genotype data. *Front. Genet.* **14**, 1175408 (2023).
32. Do, D. & Miar, Y. Evaluation of growth curve models for body weight in American mink. *Animals* **10**, 22 (2019).
33. Do, D. *et al.* PSXII-10 Development of a 70k SNP genotyping array for American mink (*Neogale vison*). *J. Anim. Sci.* **101**(Supplement\_3), 350–351 (2023).
34. Thermo Fisher Scientific, T. F. Axiom™ genotyping solution: data analysis guide. *Pub. No. 702961 Rev 5* (2017).
35. Wang, K. *et al.* PennCNV: An integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res.* **17**, 1665–1674 (2007).
36. Marioni, J. C. *et al.* Breaking the waves: Improved detection of copy number variation from microarray-based comparative genomic hybridization. *Genome Biol.* **8**, 228 (2007).
37. Kuhn, R. M., Haussler, D. & Kent, W. J. The UCSC genome browser and associated tools. *Brief. Bioinform.* **14**, 144–161 (2013).
38. Diskin, S. J. *et al.* Adjustment of genomic waves in signal intensities from whole-genome SNP genotyping platforms. *Nucleic Acids Res.* **36**, 126 (2008).
39. Garrick, D. J., Taylor, J. F. & Fernando, R. L. Deregressing estimated breeding values and weighting information for genomic regression analyses. *Genet. Select. Evol.* **41**, 1–8 (2009).
40. Lopes, F. <https://github.com/camult/DRP> (Accessed 1 Aug 2021) (2017).
41. R Core Team, A. & Team, R. C. (2022).
42. Glessner, J. T. *et al.* ParseCNV2: efficient sequencing tool for copy number variation genome-wide association studies. *Eur. J. Hum. Genet.* 1–9 (2022).
43. Chang, C. C. *et al.* Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience* **4**, 13742 (2015).
44. Glessner, J. T., Li, J. & Hakonarson, H. ParseCNV integrative copy number variation association software with quality tracking. *Nucl. Acids Res.* **41**, 64 (2013).
45. Quinlan, A. R. & Hall, I. M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
46. Montalbano, S. *et al.* Accurate and effective detection of recurrent copy number variants in large SNP genotype datasets. *Curr. Protoc.* **12**, e621 (2022).
47. Abdellaoui, A., Yengo, L., Verweij, K. J. & Visscher, P. M. 15 years of GWAS discovery: Realizing the promise. *Am. J. Hum. Genet.* **110**, 179–194 (2023).
48. Salehian-Dehkordi, H. *et al.* Genome-wide detection of copy number variations and their association with distinct phenotypes in the world's sheep. *Front. Genet.* **12**, 670582 (2021).
49. Xu, L. *et al.* Probe-based association analysis identifies several deletions associated with average daily gain in beef cattle. *BMC Genomics* **20**, 1–10 (2019).
50. Ladeira, G. C. *et al.* CNV detection and their association with growth, efficiency and carcass traits in Santa Inês sheep. *J. Anim. Breed. Genet.* **139**, 476–487 (2022).
51. Taghizadeh, S. *et al.* Genome-wide identification of copy number variation and association with fat deposition in thin and fat-tailed sheep breeds. *Sci. Rep.* **12**, 8834 (2022).
52. Qiu, Y. *et al.* Genome-wide detection of CNV regions and their potential association with growth and fatness traits in Duroc pigs. *BMC Genomics* **22**, 332 (2021).
53. Liu, M. *et al.* Diversity of copy number variation in the worldwide goat population. *Heredity* **122**, 636–646 (2019).
54. Wang, Z., Guo, Y., Liu, S. & Meng, Q. Genome-wide assessment characteristics of genes overlapping copy number variation regions in Duroc purebred population. *Front. Genet.* **12**, 753748 (2021).
55. Strillacci, M. G. *et al.* A genome-wide scan of copy number variants in three Iranian indigenous river buffaloes. *BMC Genomics* **22**, 1–14 (2021).
56. Butty, A. M. *et al.* High confidence copy number variants identified in Holstein dairy cattle from whole genome sequence and genotype array data. *Sci. Rep.* **10**, 1–13 (2020).
57. Ka, M., Chopra, D. A., Dravid, S. M. & Kim, W.-Y. Essential roles for ARID1B in dendritic arborization and spine morphology of developing pyramidal neurons. *J. Neurosci.* **36**, 2723–2742 (2016).
58. Tsurusaki, Y. *et al.* Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nat. Genet.* **44**, 376–378 (2012).

59. Celen, C. *et al.* Arid1b haploinsufficient mice reveal neuropsychiatric phenotypes and reversible causes of growth impairment. *eLife* **6**, 25730 (2017).
60. Yu, Y. *et al.* De novo mutations in ARID1B associated with both syndromic and non-syndromic short stature. *BMC Genomics* **16**, 1–10 (2015).
61. Bovo, S. *et al.* Whole-genome sequencing of European autochthonous and commercial pig breeds allows the detection of signatures of selection for adaptation of genetic resources to different breeding and production systems. *Genet. Select. Evol.* **52**, 1–19 (2020).
62. Deepa, S. S. & Dong, L. Q. APPL1: Role in adiponectin signaling and beyond. *Am. J. Physiol. Endocrinol. Metab.* **296**, 22–36 (2009).
63. Schweer, K. R. *et al.* Genome-wide association study for feed efficiency traits using SNP and haplotype models1. *J. Anim. Sci.* **96**, 2086–2098 (2018).
64. Davoudi, P. *et al.* Genome-wide detection of copy number variation in American mink using whole-genome sequencing. *BMC Genomics* **23**, 649 (2022).
65. Wilkinson, B. *et al.* TOX: An HMG box protein implicated in the regulation of thymocyte selection. *Nat. Immunol.* **3**, 272–280 (2002).
66. Seabury, C. M. *et al.* Genome-wide association study for feed efficiency and growth traits in U.S. beef cattle. *BMC Genomics* **18**, 1–25 (2017).
67. Martínez, R. In *Proceedings, 10th World Congress of Genetics Applied to Livestock Production*. 1–3.
68. De Las Heras-Saldana, S. *et al.* Use of gene expression and whole-genome sequence information to improve the accuracy of genomic prediction for carcass traits in Hanwoo cattle. *Genet. Select. Evol.* **52**, 1–6 (2020).
69. Bhuiyan, M. S. *et al.* Functional partitioning of genomic variance and genome-wide association study for carcass traits in Korean Hanwoo cattle using imputed sequence level SNP data. *Front. Genet.* **9**, 217 (2018).
70. Lee, S. H. *et al.* Genome-wide association study identifies major loci for carcass weight on BTA14 in Hanwoo (Korean cattle). *PLoS ONE* **8**, 74677 (2013).
71. Fortes, M. R. S. *et al.* Finding genes for economically important traits: Brahman cattle puberty. *Anim. Prod. Sci.* **52**, 143 (2012).
72. Li, S. *et al.* TOX and ADIPOQ gene polymorphisms are associated with antipsychotic-induced weight gain in Han Chinese. *Sci. Rep.* **7**, 45203 (2017).
73. Baranzini, S. E. *et al.* Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. *Hum. Mol. Genet.* **18**, 767–778 (2009).
74. Serão, N. V. *et al.* Single nucleotide polymorphisms and haplotypes associated with feed efficiency in beef cattle. *BMC Genet.* **14**, 94 (2013).
75. Lango Allen, H. *et al.* Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467**, 832–838 (2010).
76. Wang, K.-S. *et al.* A novel locus for body mass index on 5p15.2: A meta-analysis of two genome-wide association studies. *Gene* **500**, 80–84 (2012).
77. Gu, X. *et al.* Genome-wide association study of body weight in chicken F2 resource population. *PLoS ONE* **6**, 21872 (2011).
78. Wang, M.-S. *et al.* Comparative population genomics reveals genetic basis underlying body size of domestic chickens. *J. Mol. Cell Biol.* **8**, 542–552 (2016).
79. Zhu, M. *et al.* Polymorphisms and association of GRM1, GNAQ and HCRTR1 genes with seasonal reproduction and litter size in three sheep breeds. *Reprod. Domest. Anim.* **57**, 532–540 (2022).
80. Dong, S. *et al.* Comparative hypothalamic transcriptome analysis reveals crucial mRNAs, lncRNAs, and circRNAs affecting litter size in goats. *Genes* **14**, 444 (2023).
81. Verardo, L. *et al.* Bayesian GWAS and network analysis revealed new candidate genes for number of teats in pigs. *J. Appl. Genet.* **56**, 123–132 (2015).
82. Tahir, M. S. *et al.* Meta-analysis of heifer traits identified reproductive pathways in *Bos indicus* cattle. *Genes* **12**, 768 (2021).
83. Liu, C. *et al.* Whole-genome analysis of structural variations between Xiang pigs with larger litter sizes and those with smaller litter sizes. *Genomics* **111**, 310–319 (2019).
84. Chen, Z. *et al.* Heterozygosity and homozygosity regions affect reproductive success and the loss of reproduction: A case study with litter traits in pigs. *Comput. Struct. Biotechnol. J.* **20**, 4060–4071 (2022).
85. Krutskikh, A. *et al.* Epididymal protein Rnase10 is required for post-testicular sperm maturation and male fertility. *FASEB J.* **26**, 4198 (2012).
86. Kaloglu, C. *et al.* Wingless ligands and beta-catenin expression in the rat endometrium: The role of Wnt3 and Wnt7a/beta-catenin pathway at the embryo–uterine interface. *Mol. Reprod. Dev.* **87**, 1159–1172 (2020).
87. Chen, S. *et al.* Transcriptome analysis reveals differentially expressed genes and long non-coding RNAs associated with fecundity in sheep hypothalamus with different FecB genotypes. *Front. Cell Dev. Biol.* **9**, 633747 (2021).
88. Aloqaily, B. H., Ferranti, E. M., Summers, A. F., Gifford, C. A. & Hernandez Gifford, J. A. Intraovarian WNT3A modulates estrogen-mediated estrus behavior in cattle. *Transl. Anim. Sci.* **2**, 19–21 (2018).
89. Mattila, P. K. & Lappalainen, P. Filopodia: Molecular architecture and cellular functions. *Nat. Rev. Mol. Cell Biol.* **9**, 446–454 (2008).
90. Bohil, A. B., Robertson, B. W. & Cheney, R. E. Myosin-X is a molecular motor that functions in filopodia formation. *Proc. Natl. Acad. Sci.* **103**, 12411–12416 (2006).
91. Singh, S. K. *et al.* Melanin transfer in human skin cells is mediated by filopodia—A model for homotypic and heterotypic lysosome-related organelle transfer. *FASEB J.* **24**, 3756–3769 (2010).
92. Heimsath, E. G. Jr., Yim, Y.-I., Mustapha, M., Hammer, J. A. & Cheney, R. E. Myosin-X knockout is semi-lethal and demonstrates that myosin-X functions in neural tube closure, pigmentation, hyaloid vasculature regression, and filopodia formation. *Sci. Rep.* **7**, 17354 (2017).
93. Tokuo, H., Bhawan, J. & Coluccio, L. M. Myosin X is required for efficient melanoblast migration and melanoma initiation and metastasis. *Sci. Rep.* **8**, 10449 (2018).
94. Hirobe, T. How are proliferation and differentiation of melanocytes regulated?. *Pigment Cell Melanoma Res.* **24**, 462–478 (2011).
95. Liakath-Ali, K., Vancollie, V. E., Sequeira, I., Lelliott, C. J. & Watt, F. M. Myosin 10 is involved in murine pigmentation. *Exp. Dermatol.* **28**, 391–394 (2019).
96. Bierman, A., Guthrie, A. & Harper, C. Lavender foal syndrome in Arabian horses is caused by a single-base deletion in the MYO5A gene. *Anim. Genet.* **41**, 199–201 (2010).
97. Christen, M., de le Roi, M., Jagannathan, V., Becker, K. & Leeb, T. Myo5a frameshift variant in a miniature dachshund with coat color dilution and neurological defects resembling human griscelli syndrome type 1. *Genes* **12**, 1479 (2021).
98. Fontanesi, L., Scotti, E., Dall’Olio, S., Oulmouden, A. & Russo, V. Identification and analysis of single nucleotide polymorphisms in the myosin VA (MYO5A) gene and its exclusion as the causative gene of the dilute coat colour locus in rabbit. *World Rabbit Sci.* **20**, 35–41 (2012).
99. Manakhov, A. D., Andreeva, T. V., Trapezov, O. V., Kolchanov, N. A. & Rogaev, E. I. Genome analysis identifies the mutant genes for common industrial Silverblue and Hedlund white coat colours in American mink. *Sci. Rep.* **9**, 4581 (2019).

100. Kovalevich, J., Tracy, B. & Langford, D. PINCH: More than just an adaptor protein in cellular response. *J. Cell. Physiol.* **226**, 940–947 (2011).
101. Legate, K. R., Montañez, E., Kudlacek, O. & Füssler, R. ILK, PINCH and parvin: The tIPP of integrin signalling. *Nat. Rev. Mol. Cell Biol.* **7**, 20–31 (2006).
102. Lorenz, K. *et al.* Integrin-linked kinase is required for epidermal and hair follicle morphogenesis. *J. Cell Biol.* **177**, 501–513 (2007).
103. Endo, C. *et al.* Genome-wide association study in Japanese females identifies fifteen novel skin-related trait associations. *Sci. Rep.* **8**, 8974 (2018).
104. Karaköse, E. *et al.* The focal adhesion protein PINCH-1 associates with EPLIN at integrin adhesion sites. *J. Cell Sci.* **128**, 1023–1033 (2015).
105. Jepsen, J. R. *et al.* Aleutian mink disease virus and humans. *Emerg. Infect. Dis.* **15**, 2040 (2009).
106. Regis, G., Conti, L., Boselli, D. & Novelli, F. IFN $\gamma$ R2 trafficking tunes IFN $\gamma$ –STAT1 signaling in T lymphocytes. *Trends Immunol.* **27**, 96–101 (2006).
107. Schroder, K., Hertzog, P. J., Ravasi, T. & Hume, D. A. Interferon- $\gamma$ : An overview of signals, mechanisms and functions. *J. Leukoc. Biol.* **75**, 163–189 (2004).
108. Holzer, U., Reinhardt, K., Lang, P., Handgretinger, R. & Fischer, N. Influence of a mutation in IFN- $\gamma$  receptor 2 (IFNGR2) in human cells on the generation of Th17 cells in memory T cells. *Hum. Immunol.* **74**, 693–700 (2013).
109. Akhter, N. *et al.* Apurinic/aprimidinic endonuclease 1/redox factor-1 (Ape1/Ref-1) modulates antigen presenting cell-mediated T helper cell type 1 responses. *J. Biol. Chem.* **291**, 23672–23680 (2016).
110. Oliveira, T. T. *et al.* APE1/Ref-1 role in inflammation and immune response. *Front. Immunol.* **13**, 726 (2022).
111. Furumai, R., Tamada, K., Liu, X. & Takumi, T. UBE3A regulates the transcription of IRE, an antiviral immunity. *Hum. Mol. Genet.* **28**, 1947–1958 (2019).
112. Zhang, N. *et al.* UBE3A deletion enhances the efficiency of immunotherapy in non-small-cell lung cancer. *Bioengineered* **13**, 11577–11592 (2022).
113. Prekeris, R., Klumperman, J. & Scheller, R. H. Syntaxin 11 is an atypical SNARE abundant in the immune system. *Eur. J. Cell Biol.* **79**, 771–780 (2000).
114. D'Orlando, O. *et al.* Syntaxin 11 is required for NK and CD8+ T-cell cytotoxicity and neutrophil degranulation. *Eur. J. Immunol.* **43**, 194–208 (2013).

## Acknowledgements

We are very grateful to the Canadian Center for Fur Animal Research, Millbank Fur Farm, North American Fur Auction, SAGA, and Miara Lab staff for collecting and providing the data.

## Author contributions

M.S., G.P., Z.W. and Y.M.: conceived and designed the experiments. P.D.: performed the experiments and analyzed the data. P.D., D.D., B.R., S.C., and Y.M.: interpreted the results. P.D.: wrote the main manuscript. D.D., G.H. and S.V.: collaborated in data preparations and estimated the deregressed EBVs. Y.M.: supervised the project. D.D., B.R., S.C., G.P. and Y.M. reviewed and revised the manuscript. M.S., G.P., Z.W. and Y.M. acquired the financial support for the project. All authors have read and agreed to the published version of the manuscript.

## Funding

Financial support was received from the Natural Sciences and Engineering Research Council of Canada (NSERC), Mitacs, Canada Mink Breeders Association, Nova Scotia Mink Breeders Association, Nova Scotia Department of Agriculture, and Mink Veterinary Consulting and Research Service Ltd. This research was enabled in part by support provided by ACENET (<http://www.ace-net.ca>) and the Digital Research Alliance of Canada (<http://www.alliancecan.ca>).

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-50497-3>.

**Correspondence** and requests for materials should be addressed to Y.M.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024