Diverging transposon activity among polar bear sub-populations inhabiting different climate zones

Authors:

Alice M. Godden[1], Benjamin T. Rix[1] & Simone Immler[1]

Affiliations:

[1] School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, United Kingdom

Corresponding author: Alice M. Godden, Email: alice.godden@uea.ac.uk

Keywords: Transposable elements, thermal stress, adaptive response, Ursus, polar bear

Abstract

A new subpopulation of polar bears (*Ursus maritimus*) was recently discovered in the South-East of Greenland (SEG). This isolated colony inhabits a warmer climate zone, akin to the predicted future environments of polar bears with vastly reduced sea ice habitats, rendering this population of bears particularly important. Over two-thirds of polar bears will be extinct by 2050 with total extinction predicted by the end of this century and understanding possible mechanisms of adaptation via genomic analyses and preservation are critical. Transposable elements (TEs) are parasitic mobile elements that may play a role in an adaptive response to environmental challenges. We analysed transcriptome data from polar bear sub-populations in North-East and South-East Greenland (NEG, SEG), who reside in cooler and warmer habitats respectively, to identify differentially expressed, divergent TE species, TE families and linked changes in gene expression with some overlapping significantly differentially expressed TEs and genes. We identified activity hotspots in the genome of regions with significantly differentially expressed TEs. LINE family TEs were the most abundant, and most differentially expressed and divergent in the SEG population compared to reference TEs. Our results provide insights into how a genomic response at the TE level may allow the SEG subpopulations adapt and survive to climate change and provides a useful resource for conservation in polar bears.

Abbreviations:

DMI - Danish Meteorological Institute

DNA-seq – DNA sequencing

lncRNAs – Long non-coding RNAs

NEG North-East Greenland

piRNA- piwi-interacting RNA

RNA-seq – RNA sequencing

SEG – South-East Greenland

TEs – Transposable elements

Background

Global average temperatures have now risen to more than 1°C above the pre-industrial era (1). The Arctic current is the warmest it has been in the last 125,000 years and temperatures continue to rise (2, 3). In the South East of Greenland (SEG), the ice sheet margin is rapidly receding, causing vast ice and habitat loss (4). North-East Greenland (NEG) is an Arctic tundra, whereas SEG is covered by forest tundra (5). The SEG climate has high levels of precipitation, wind, and steep coastal mountains (6). This challenging habitat has led to severe warnings for the survival of the polar bear *Ursus maritimus*, with the probability of a reduction in population size by over 90% within 40 years set at 0.71. The IUCN have listed the polar bear as a vulnerable species (7), and understanding possible ways of adaptation to climate change in such an enigmatic animal at the genome level is key.

Environmental stress is reported to have a significant impact on genome structure, including the expression and mobilisation of transposable elements (TEs), (8). TEs are mobile genetic elements that can shape evolution through self-replication and insertion events in the genome, which in turn may change gene expression patterns (9, 10). TEs inserted into or near a promoter region, also alter the expression of genes and generate novel regulatory networks (11). TEs make up 38.1% of the genome in polar bears (12), which is similar to the giant panda *Ailuropoda melanoleuca* at 37%, and 36.1% of the dog *Canis familiaris* genome (13), and hence TEs may play an important part in the process of adaptation. Current data shows that over 150,000 genome insertion mutations have been generated by TEs, driving variation in ursine and tremarctine bears (12). The aim of this study was to investigate the transcriptomes of NEG and SEG polar bears and test for TE divergence causing genetic differences and variation in gene expression between these sub-populations because of the rising temperatures.

Methods

Meteorological analysis

We collected and analysed meteorological data on the climate in NEG and SEG from the Danish Meteorological Institute (14). SEG locations included: Aputiteeq, Tasiilaq, Mittarfik, Kulusuk, Ikermiit and Qaqortoq. NEG locations included: Ittoqqortoormiit, Daneborg, Port of Denmark and Station Nord. Temperature data covering the highest, and lowest recorded temperatures each year from 1958-2024 were included and the data was plotted with Tableau Public version 2024.1 (See Supplementary File 1 for latitude and longitude co-ordinates).

RNA-seq analysis

Transcriptomic data from blood samples were processed bioinformatically with published reference polar bear genome ASM1731132v1. Transcriptomic data from blood samples of adult polar bears were downloaded from the European Nucleotide Archive under accession no. PRJNA669153 (15) (see Suppl. File 2 for sample accession numbers and metadata). For RNA-sequencing (RNA-seq), total RNA was used, with haemoglobin depletion by CRISPR-Cas9 to generate libraries, which were sequenced on Illumina HiSeqX to generate 151 bp paired-end reads (15). We selected samples that had similar coverage depth resulting in samples from five female NEG and seven female SEG bears and two male NEG and three male SEG bears. NF-core RNA-Seq pipeline version 1.4.2 was used to analyse the RNA-seq data (https://nf-co.re/rnaseq/1.4.2/) (16). Genome and annotation files were used and accessed as follows: ASM1731132v1 (GCF_017311325.1) GCF_017311325.1_ASM1731132v1_genomic.fna, and annotation GCF_017311325.1_ASM1731132v1_genomic.gtf https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_017311325.1/). Ursus_maritimus.UrsMar_1.0.110.gtf.gz, (https://ftp.ensembl.org/pub/release-110/gtf/ursus_maritimus/), Ursus_maritimus.UrsMar_1.0.dna.nonchromosomal.fa.gz, (https://ftp.ensembl.org/pub/release-110/fasta/ursus_maritimus/dna/). ASM1731132v1 is the most recent genome assembly and was used here as the reference genome.

We used *DESeq2* (17) in R (18) to run a differential expression analyses between SEG and NEG bears. The model used in *DESeq2* to generate a dds object was *design = ~ sex + location* (location referring to NEG or SEG). This was to account for the unequal sample sizes for each sex and to focus on the differences in gene and TE expression from different geographical locations. To link the impact of temperature to the transcriptome data, the average temperature over the year preceding blood sample collection recorded in the town nearest to the bear's location was included in the analyses (mean temperature over 2016, bears sampled in 2017). The following model design was used in *DESeq2* analysis: *design = ~ sex + mean_temp_2016 + location.* Results were then extracted as follows: *res_temperature <- results(dds, name="mean_temp_2016")* to examine main impact of temperature on differential expression of genes and TEs. Because location and temperature were colinear, the additive model was used without an interaction term to avoid redundancy.

The significance cutoff for differentially expressed genes and TEs was $\log_2$ fold change > 1 and a p_{adj} > 0.05. Significantly differentially expressed genes were fed into *ShinyGO* v0.77 (19), with a background list of all genes found to be expressed in this study, to conduct GO terms analyses, and results were plotted using custom Python scripts.

TE analysis

We used *Telescope* v1.0.3 to identify differentially expressed TE species (20). *RepeatMasker* (21) 4.1.1 and *RepeatModeler* 2.0.3 (22) were used to assess Kimura distances. We merged the bam files of individual samples by treatment group, then indexed and converted to fasta format with *Samtools* (23). *RepeatMasker* was used as follows: *RepeatMasker -species 29073 -par 5 -s -a -nolow -no_is*, with fasta inputs split into 200 chunks for parallelisation. Aligned outputs were merged and calcDivergenceFromAlign.pl was used for analysis. Genome sizes in the converted fasta files were calculated using calcGenomesize.pl. This version of RepeatMasker utilised Dfam 3.2, with genome version UrsMar1.0 TE annotations. TEs are tissue specific in expression and activity (24), and the RepeatMasker default library for TEs in polar bears utilizes an UrsMar1.0 based annotation,

therefore a custom library, "-lib", was generated for the ASM1731132v1 reference genome assembly. To generate a custom library "-lib" for use with RepeatMasker with the GCF_017311325.1_ASM1731132v1_genomic.fna as reference genome for the ASM1731132v1 assembly, as the RepeatMasker defaults to an UrsMar1.0 annotation. To create a RepeatModeler database*: BuildDatabase -name custom_repeats_db_ASM1731132v1 -engine ncbi "$GENOME_FILE"*. Followed by: *RepeatModeler -database custom_repeats_db_ASM1731132v1 -pa 40 -engine ncbi*. The output was used as the -lib input instead of species in the RepeatMasker pipeline. RepeatModeler 2.0.3, ncbi-blast 2.10.1 and rmblast 2.10.0 (25), were used to analyse TE divergence. To produce the TE gtf annotation file for Telescope pipeline, the perl scripts from TETranscripts (26) were used as follows: *perl makeTEgtf.pl -c 1 -s 2 -e 3 -o 6 -t 4 ASM1731132v1_rmsk > ASM1731132v1_rmsk_TE.gtf* and likewise for UrsMar1.0.

Data analysis

All R scripts and Python scripts used to generate figures can be found at: https://github.com/alicegodden/polarbear/. Rstudio version 4.1.1 (27), and the following packages were used in the analysis and presentation of the data in this project. These included: *DESeq2* 1.34.1 (17), *Tidyverse* 2.0.0 (28), *EnhancedVolcano* 1.12.0 (29), *ggplot2* 3.4.2 (30), *Readr* 2.1.4 (31), *tidyr* v1.3.0 (32), *viridis* v0.6.4 (33), *reshape* (34), *hrbrthemes* (35) and *gridExtra* (36).

Results

Meteorological analysis

The meteorological observation data from Danish Meteorological Institute (DMI), revealed vast variation in average lowest temperature measured from 1958-2024 between NEG and SEG (Fig. 1A), as denoted by the 95th percentile size of dots representing more variation. The variance in observed lowest average temperatures was higher in SEG compared to NEG (Fig. 1B). Tasiilaq (SEG) showed the highest reported temperature, and Station Nord (NEG) showed the lowest recorded temperatures (Fig. 1A).

Gene & TE expression analysis

We initially identified 179 significantly differentially expressed TEs ($P_{adj} \leq 0.05$) between SEG and NEG bears when including both sex and location into the model (Fig. Suppl. Fig. 1). A Principal Components Analysis showed modest clustering of samples based on geographical location (Suppl. Fig. 1A). The majority of the significantly differentially expressed TE species belong to the LINE family (Suppl. Fig. 1B-C). The most significantly downregulated TE was DNA TE hAT-Charlie_dup1276 from the Charlie 8 subfamily. The most significantly enriched TE was LINE TE L1_dup3552 from the L1 Carn1 subfamily (Suppl. Fig. 1B). To find genomic regions where TE activity may originate from, the genomic positions of significantly differentially expressed TE species were plotted (Suppl. Fig. 1D), revealing some clustering of differentially expressed TEs on chromosomes NW_024423453.1, NW_024424119.1 and NW_02442563.1. No significantly differentially expressed TE species were identified on the Y chromosome (NW_024423336.1), but a few LINE TE species were located on the X chromosome (NW_024423319.1).

Clustering analysis showed some grouping of samples on PC1 with up to 29% variance and 15% for TEs based on the geographical origin and local temperature of the sample (Suppl. Fig. 2A). Differential gene expression analysis for *sex* + *location* as covariates revealed 13 significantly differentially expressed genes (Suppl. Fig. 2B). RSAD2 was most significantly downregulated and LOC10367587 was most significantly enriched in SEG v NEG bears. Of note, HSPH1 and HSPA4L were significantly enriched in SEG bears (Suppl. Fig. 2B). GO terms analyses of biological processes, molecular function, and cellular function for the significantly differentially expressed genes from ASM1731132v1 alignments with *sex* + *location* covariates highlighted enrichment for the following terms: internal cellular dynamics and movements: chromosome movement towards spindle pole, PRC1 complex, organic acid:sodium symporter activity, negative regulation of protein localization to nucleus, and immune responses: negative regulation of viral genome replication, defence response to virus, defence response to symbiont, (Suppl. Fig. 2C).

When adding temperature as a covariate into our analyses samples clustered based on latitude and mean temperature of the region where sampled (Fig. 2B). We also identified 1,534 significantly differentially expressed TEs (Fig. 2). The most significantly enriched TE was L2b LINE L2_dup45185, and the most significantly downregulated TE was L2a LINE/L2 dup_44480 (Fig. 2A). The most abundant was again the LINE class TEs (Fig. 2C). Age estimations of differentially expressed TEs showed divergence in TEs between the SEG and NEG bear populations (Fig. 2 D-E). With both populations displaying divergence most strongly in LINE family TEs, then DNA and LTR TEs. In the SEG population (Fig. 2E), we saw an overall increase in TE divergence and a notable peak of younger divergent LINE TEs.

In our gene expression analysis, as with the TEs, clustering analysis showed some clustering of polar bear samples based on latitude and temperature of the geographic area when sampled (Fig. 3A). We identified 27 significantly differentially expressed genes, with LOC103677657 most downregulated and LOC103668192 most enriched (Fig. 3B). We found no overlap in position of significantly differentially expressed genes and TEs, but when using sex + location covariates, but addition of *temperature* covariate for ASM1731132v1 revealed 10 overlapping genes with TEs (Fig. 3D). These genes were: LOC121104823, LOC103667593, LOC103678552, LOC103677657, LOC121100717, LOC121101012, BBS12 (ENSUMAG00000004608- Bardet-Biedl syndrome 12), LOC121104992, SHANK1 and LOC103660844.

Discussion

Impact of rising and varied temperature

Polar bears are dependent on sea ice, and the warmer and more variable SEG climate (Fig. 1A) is forecast to be the norm by the end of the 21st century (15). The warmest temperatures in this study were observed in Tasiilaq (Fig. 1B), which is characterised by a deep fjord and mountainous habitat with some drifting sea pack ice as evidence of how temperature is isolating of the SEG bears (6,

37). We therefore assessed how variation in temperature was associated with the expression of TEs and genes.

TE expression

Environmental stressors may affect the activity and expression levels of the mobile and non-coding genome and lead to genomic adaptation (8). Positive effects of TE mobilisation include beneficial mutations accelerating adaptation (38), and the generation of novel piRNA clusters that can act in a feedback loop to repress further TE action and invasion on the host genome (8). In contrast, negative effects may include disruption to gene expression through loss of function mutations, as TEs are found to insert themselves predominantly into genome regions that allows for their successful propagation, although these are extremely rare (9). TEs can therefore lead to the generation of novel networks by assimilating into the promoter regions of other genes, thus perturbing expression of themselves and up to hundreds of other genes (11, 39, 40).

The main clusters of TEs identified in our data belong to the LINE class, these were also the most abundant and significantly differentially expressed in the SEG bears (Fig. 2). LINE and SINE TEs are the most abundant family of TEs in mammalian genomes, with limited LTR and DNA TE retention (11). LINE and SINE TEs are also the most mobile and active TEs in bear genomes, and are most active in brown bears, followed by polar bears (12). To see if climate was impacting their expression, we included impact of temperature in our analysis. This identified 1,534 significantly differentially expressed TEs (Fig. 2A, C). This linking of temperature with genome mobility in polar bear aligns with similar studies research in model systems including *Arabidopsis* (41), *C. elegans* (42) and *Drosophila* (43), where environmental stress was reported to increase TE mobility. Future work could include the use of DNA-sequencing (DNA-seq) and small RNA-seq data to analyse for TE insertion events and differential expression analysis of piRNAs.

Our findings showed a putative role for TEs in the adaptation of SEG polar bears by TE divergence. This offers new avenues to understanding rapid adaptation to changing environmental conditions. Divergent TEs between populations have also been described in great pandas, where LINE and SINE TEs were generally younger in age, indicative of recent mobilisation events (13). Our study also found an abundance of LINE TEs for both the NEG and SEG bear populations, potentially highlighted recent genome expansion of RNA class 1 TEs (Fig. 2C). Additionally, LINE TEs have been previously shown to be important in the evolution of regulatory elements in the mammalian genome (44). This is significant as we observed a peak of younger, divergent LINE TEs in the SEG population of polar bears (Fig. 2D-E). Older TEs included a modest amount of DNA and LTR TEs in older genome expansion events, although these divergent events constitute a small percentage of the genome (Fig. 2D-E). This expansion LINE TEs observed in SEG bears could be in response to their challenging habitat and climate.

Gene expression

The evolutionary origins of bears as a group are complicated, having evolved over the past 5 million years in harsh and changing environments (45). Polar and brown bears *Ursus arctos* are the most closely related, with 8.8% of brown bear genomes originating from polar bears (45, 46). To better understand polar bear genomes, we therefore profiled the differential gene expression of SEG versus NEG polar bears, whilst using *sex* as a covariate to account for differences in sample numbers and sex biases in gene expression. Having a more balanced dataset with more male samples, and more NEG bears in the future will help investigate these differences further. Given the differences in habitat over the recent decades of extreme warming and variable weather patterns in SEG, it is unsurprising that SEG bears are starting to show significant differences in gene expression compared to NEG bears (6).

When analysing sex and location as covariates, genes that were differentially expressed in SEG bears included: RSAD2 (downregulated), ACTR2, HSPA4L and HSPH1 (enriched) (Suppl. Fig. 2B). RSAD2 (also known as Viperin) is an evolutionarily conserved antiviral protein-coding gene

(47). ACTR2 (actin-related protein 2) in house mice *Mus musculus* is involved in thrombosis, and mutations in the gene have been found to be linked to thrombo-suppression (48). HSPA4L (Hsp70) and HSPH1 (Hsp110) are highly conserved and related families of proteins known as heat-shock proteins that have been found in all eukaryotic species so far, acting as molecular chaperones (49), with other Hsp chaperones, including Hsp90, to regulate protein folding, unfolding and degradation (50). Hsp90 and Hsp70 are also important in piRNA biogenesis and silencing of TEs, and function in elevated temperature environments (51). The most significant differential expression in TEs was observed in the analysis with *temperature* as a covariate (Fig. 2A). This is supported by evidence in *Drosophila* where TEs are activated in response to heat shock due to Hsp70 and other chaperones having a role in the collapse of piRNA production, and thus further TE action (52). Combined, this is evidence that the host genome is potentially responding to the changes in environmental temperature.

We tried to further link gene expression with TE expression in the differentially expressed genes when taking temperature as covariate into account. We observed 10 genes that were significantly differentially expressed and overlapped with significantly differentially expressed TEs, these included: LOC121104823, LOC103678552, LOC121100717, and LOC121104992 (all enriched), LOC121101012 (downregulated) are all long non-coding RNAs (lncRNAs) (Fig. 3B, D). Long non-coding RNAs are highly abundant in mammals, with over 80% of transcribed products generating lncRNAs that have roles in chromatin organization and gene regulation (53). The accessibility of chromatin, and methylation is significant as this can lead to de-repression of TEs (54).

Other genes of interest included LOC103677657 a pseudogene with unknown function, LOC103667593 which is a predicted mRNA for filaggrin-2-like protein linked to inflammatory skin diseases in human (55), and LOC121104992, a predicted mRNA for keratin-associated 4-8-like protein. BBS12 was modestly downregulated in the SEG bears and is an mRNA for Bardet-Biedl syndrome 12 protein, involved in chaperone protein complex assemblies, and is orthologous with human *BBS12*. Mutations or loss of BBS12 in humans leads to cilia abnormalities (56). SHANK1 was upregulated in SEG bears and is an mRNA for SH3 and multiple ankyrin repeat domains

protein 1 and orthologous to human *SHANK1*. *SHANK1* encodes a scaffold protein and can inhibit the activation of integrins (57), which can lead to changes in cell signalling. This links to the GO term processes observed for phosphatidylinositol activity highlighting cellular activity and signalling (Fig. 3C). Other GO terms in response to the warmer SEG climate showed enrichment for terms related to immune responses including defence response to virus, and negative regulation of viral genome replication (Suppl. Fig. 2C). TE invasion is known to induce an immune response in mammals, activating the interferon and innate immune pathways (58). Polar bears in regions of reduced sea ice habitat are at higher risk of pathogen exposure (59), which may explain the changes in SEG bears. However, whether SEG bears are indeed exposed to higher pathogen loads needs further careful investigations.

Conclusions

Our work revealed key differences in the genomes of SEG bears compared to the NEG bears, most strikingly in the mobile genome, potentially because of more variable and higher temperatures in SEG. The main differences we found were the significantly differentially expressed TEs and divergent genome expansion events exhibited by LINEs. Our study provides a starting point for further work into TEs in a wider subset of polar bear subpopulations and highlights potential candidate TE species for further investigation as potential markers of environmental stress. Future work to revise and further annotate the *Ursus maritimus* genome, at all levels not just TEs will prove highly valuable in the efforts to preserve this vulnerable species through genomic analyses. Larger scale analysis across all other subpopulations of polar bear, and comparative analyses across different bear species in a range of climate zones may prove highly valuable for the planning of conservation efforts and species management.

References

1. Masson-Delmotte V, Zhai P, Pörtner H-O, Roberts D, Skea J, Shukla PR. Global Warming of 1.5 C: IPCC special report on impacts of global warming of 1.5 C above pre-industrial levels in context of strengthening response to climate change, sustainable development, and efforts to eradicate poverty: Cambridge University Press; 2022.
2. Stendel M, Christensen JH, Petersen D. Arctic climate and climate change with a focus on Greenland. Advances in Ecological Research. 2008;40:13-43.
3. Horhold M, Munch T, Weissbach S, Kipfstuhl S, Freitag J, Sasgen I, et al. Modern temperatures in central-north Greenland warmest in past millennium. Nature. 2023;613(7944):503-7.
4. Bjørk AA, Kjær KH, Korsgaard NJ, Khan SA, Kjeldsen KK, Andresen CS, et al. An aerial view of 80 years of climate-related glacier fluctuations in southeast Greenland. Nature Geoscience. 2012;5(6):427-32.
5. Normand S, Randin C, Ohlemuller R, Bay C, Hoye TT, Kjaer ED, et al. A greener Greenland? Climatic potential and long-term constraints on future expansions of trees and shrubs. Philos Trans R Soc Lond B Biol Sci. 2013;368(1624):20120479.
6. Berdahl M, Rennermalm A, Hammann A, Mioduszweski J, Hameed S, Tedesco M, et al. Southeast Greenland winter precipitation strongly linked to the Icelandic low position. Journal of Climate. 2018;31(11):4483-500.
7. Regehr EV, Laidre KL, Akcakaya HR, Amstrup SC, Atwood TC, Lunn NJ, et al. Conservation status of polar bears (Ursus maritimus) in relation to projected sea-ice declines. Biol Lett. 2016;12(12).
8. Godden AM, Immler S. The potential role of the mobile and non-coding genomes in adaptive response. Trends Genet. 2022.
9. Bourque G, Burns KH, Gehring M, Gorbunova V, Seluanov A, Hammell M, et al. Ten things you should know about transposable elements. Genome Biol. 2018;19(1):199.
10. Chang NC, Rovira Q, Wells J, Feschotte C, Vaquerizas JM. Zebrafish transposable elements show extensive diversification in age, genomic distribution, and developmental expression. Genome Res. 2022;32(7):1408-23.
11. Platt RN, 2nd, Vandewege MW, Ray DA. Mammalian transposable elements and their impacts on genome evolution. Chromosome Res. 2018;26(1-2):25-43.
12. Lammers F, Gallus S, Janke A, Nilsson MA. Phylogenetic Conflict in Bears Identified by Automated Discovery of Transposable Element Insertions in Low-Coverage Genomes. Genome Biol Evol. 2017;9(10):2862-78.
13. Peng C, Niu L, Deng J, Yu J, Zhang X, Zhou C, et al. Can-SINE dynamics in the giant panda and three other Caniformia genomes. Mob DNA. 2018;9:32.
14. Weather archive [Internet]. Danish Meteorological Institute. 2024. Available from: Danish Meteorological Institute.
15. Laidre KL, Supple MA, Born EW, Regehr EV, Wiig O, Ugarte F, et al. Glacial ice supports a distinct and undocumented polar bear subpopulation persisting in late 21st-century sea-ice conditions. Science. 2022;376(6599):1333-8.
16. Ewels PA, Peltzer A, Fillinger S, Patel H, Alneberg J, Wilm A, et al. The nf-core framework for community-curated bioinformatics pipelines. Nat Biotechnol. 2020;38(3):276-8.
17. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.
18. Team RC. R: A Language and Environment for Statistical Computing R Foundation for Statistical Computing, Vienna, Austria. 2017.
19. Ge SX, Jung D, Yao R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. Bioinformatics. 2020;36(8):2628-9.
20. Bendall ML, de Mulder M, Iniguez LP, Lecanda-Sanchez A, Perez-Losada M, Ostrowski MA, et al. Telescope: Characterization of the retrotranscriptome by accurate estimation of transposable element expression. PLoS Comput Biol. 2019;15(9):e1006453.
21. Smit AF, Hubley R, Green P. RepeatMasker software 2013.
22. Smit AF, Hubley R. RepeatModeler Open-1.0. 2015.
23. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754-60.
24. Karttunen K, Patel D, Xia J, Fei L, Palin K, Aaltonen L, et al. Transposable elements as tissue-specific enhancers in cancers of endodermal lineage. Nat Commun. 2023;14(1):5313.
25. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10:421.
26. Jin Y, Tam OH, Paniagua E, Hammell M. TEtranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. Bioinformatics. 2015;31(22):3593-9.
27. Rstudio. RStudio: Integrated Development Environment for R. 2021:http://www.rstudio.com/.
28. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, et al. Welcome to the tidyverse. Journal of Open Source Software. 2019;4(43):1-6.
29. Blighe K, Rana S, Lewis M. EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling
. 2021:https://github.com/kevinblighe/EnhancedVolcano.
30. Ginestet C. ggplot2: Elegant Graphics for Data Analysis. J Roy Stat Soc A. 2011;174:245-.
31. Wickham H, Hester, J., Bryan, J.,. readr: Read Rectangular Text Data. 2023: https://CRAN.R-project.org/package=readr.
32. Wickham H VD, Girlich M,. tidyr: Tidy Messy Data. R package version 1.3.0. 2023:https://github.com/tidyverse/tidyr.

33. Garnier S, Ross N, Rudis B, Sciaini M, Scherer C. viridis(Lite) - Colorblind-Friendly Color Maps for R. 2023:10.5281/zenodo.4679423.
34. Wickham H. reshape: Flexibly Reshape Data. 2022.
35. Rudis B, Kennedy P, Reiner P, Wilson D, Adam X, Barnett J, et al. hrbrthemes: Additional Themes, Theme Components and Utilities for 'ggplot2'. 2024.
36. Auguie B, Antonov A. gridExtra: Miscellaneous Functions for "Grid" Graphics. cran. 2017.
37. Wiig Ø, Henrichsen P, Sjøvold T, Born EW, Laidre KL, Dietz R, et al. Variation in non-metrical skull traits of polar bears (Ursus maritimus) and relationships across East Greenland and adjacent subpopulations (1830–2013). Polar Biology. 2019;42:461-74.
38. Quadrana L, Etcheverry M, Gilly A, Caillieux E, Madoui MA, Guy J, et al. Transposition favors the generation of large effect mutations that may facilitate rapid adaption. Nat Commun. 2019;10(1):3421.
39. Chuong EB, Elde NC, Feschotte C. Regulatory activities of transposable elements: from conflicts to benefits. Nat Rev Genet. 2017;18(2):71-86.
40. Fueyo R, Judd J, Feschotte C, Wysocka J. Roles of transposable elements in the regulation of mammalian transcription. Nat Rev Mol Cell Biol. 2022;23(7):481-97.
41. Roquis D, Robertson M, Yu L, Thieme M, Julkowska M, Bucher E. Genomic impact of stress-induced transposable element mobility in Arabidopsis. Nucleic Acids Res. 2021;49(18):10431-47.
42. Garrigues JM, Tsu BV, Daugherty MD, Pasquinelli AE. Diversification of the Caenorhabditis heat shock response by Helitron transposable elements. Elife. 2019;8.
43. Garcia Guerreiro MP. What makes transposable elements move in the Drosophila genome? Heredity (Edinb). 2012;108(5):461-8.
44. Roller M, Stamper E, Villar D, Izuogu O, Martin F, Redmond AM, et al. LINE retrotransposons characterize mammalian tissue-specific and evolutionarily dynamic regulatory regions. Genome Biol. 2021;22(1):62.
45. Kumar V, Lammers F, Bidon T, Pfenninger M, Kolter L, Nilsson MA, et al. The evolutionary history of bears is characterized by gene flow across species. Sci Rep. 2017;7:46487.
46. Cahill JA, Stirling I, Kistler L, Salamzade R, Ersmark E, Fulton TL, et al. Genomic evidence of geographically widespread effect of gene flow from polar bears into brown bears. Mol Ecol. 2015;24(6):1205-17.
47. Upadhyay AS, Vonderstein K, Pichlmair A, Stehling O, Bennett KL, Dobler G, et al. Viperin is an iron-sulfur protein that inhibits genome synthesis of tick-borne encephalitis virus via radical SAM domain activity. Cell Microbiol. 2014;16(6):834-48.
48. Westrick RJ, Tomberg K, Siebert AE, Zhu G, Winn ME, Dobies SL, et al. Sensitized mutagenesis screen in Factor V Leiden mice identifies thrombosis suppressor loci. Proc Natl Acad Sci U S A. 2017;114(36):9659-64.
49. Easton DP, Kaneko Y, Subjeck JR. The hsp110 and Grp1 70 stress proteins: newly recognized relatives of the Hsp70s. Cell Stress Chaperones. 2000;5(4):276-90.
50. Rosenzweig R, Nillegoda NB, Mayer MP, Bukau B. The Hsp70 chaperone network. Nat Rev Mol Cell Biol. 2019;20(11):665-80.
51. Karam JA, Parikh RY, Nayak D, Rosenkranz D, Gangaraju VK. Co-chaperone Hsp70/Hsp90-organizing protein (Hop) is required for transposon silencing and Piwi-interacting RNA (piRNA) biogenesis. J Biol Chem. 2017;292(15):6039-46.
52. Cappucci U, Noro F, Casale AM, Fanti L, Berloco M, Alagia AA, et al. The Hsp70 chaperone is a major player in stress-induced transposable element activation. Proc Natl Acad Sci U S A. 2019;116(36):17943-50.
53. Karlsson O, Baccarelli AA. Environmental Health and Long Non-coding RNAs. Curr Environ Health Rep. 2016;3(3):178-87.
54. Zhou W, Liang G, Molloy PL, Jones PA. DNA methylation enables transposable element-driven genome expansion. Proc Natl Acad Sci U S A. 2020;117(32):19359-66.
55. Zeitvogel J, Jokmin N, Rieker S, Klug I, Brandenberger C, Werfel T. GATA3 regulates FLG and FLG2 expression in human primary keratinocytes. Sci Rep. 2017;7(1):11847.
56. Seo S, Baye LM, Schulz NP, Beck JS, Zhang Q, Slusarski DC, et al. BBS6, BBS10, and BBS12 form a complex with CCT/TRiC family chaperonins and mediate BBSome assembly. Proc Natl Acad Sci U S A. 2010;107(4):1488-93.
57. Lilja J, Zacharchenko T, Georgiadou M, Jacquemet G, De Franceschi N, Peuhu E, et al. SHANK proteins limit integrin activation by directly interacting with Rap1 and R-Ras. Nat Cell Biol. 2017;19(4):292-305.
58. Gazquez-Gutierrez A, Witteveldt J, S RH, Macias S. Sensing of transposable elements by the antiviral innate immune system. RNA. 2021;27(7):735-52.
59. Rode KD, Van Hemert C, Wilson RR, Woodruff SP, Pabilonia K, Ballweber L, et al. Increased pathogen exposure of a marine apex predator over three decades. PLoS One. 2024;19(10):e0310973.

Acknowledgements

We thank Dr. Oliver Tam (Hammel lab), for TE annotation file generation support. We also thank Jean-Charles de Coriolis for continuing and engaging TE discussions. This study was funded by

grants from the Natural Environment Research Council (NE/S011188/1) and the European Research Council (SELECTHAPLOID - 101001341) to SI.

Declarations

Authors declare no conflict or competing interests.

Author contributions

AMG conceived the idea and conducted the bioinformatic analyses, and wrote the manuscript. BR contributed to data analysis and conducted metereological analyses. SI supported manuscript preparation and funding.

Data and code availability

RNA-seq data was deposited by (15) under ENA accession PRJNA669153. All R and python scripts used to analyse this data can be found in the following Github repository: https://github.com/alicegodden/polarbear/ . Metadata used for the files accessed can be found in Supplementary file 2. The datasets supporting this article have been uploaded as part of the supplementary material in files 1-8. Supplementary files include all raw counts data from RNA-seq analyses, DESeq2 outputs for differential gene and TE expression with different models.

FIGURES

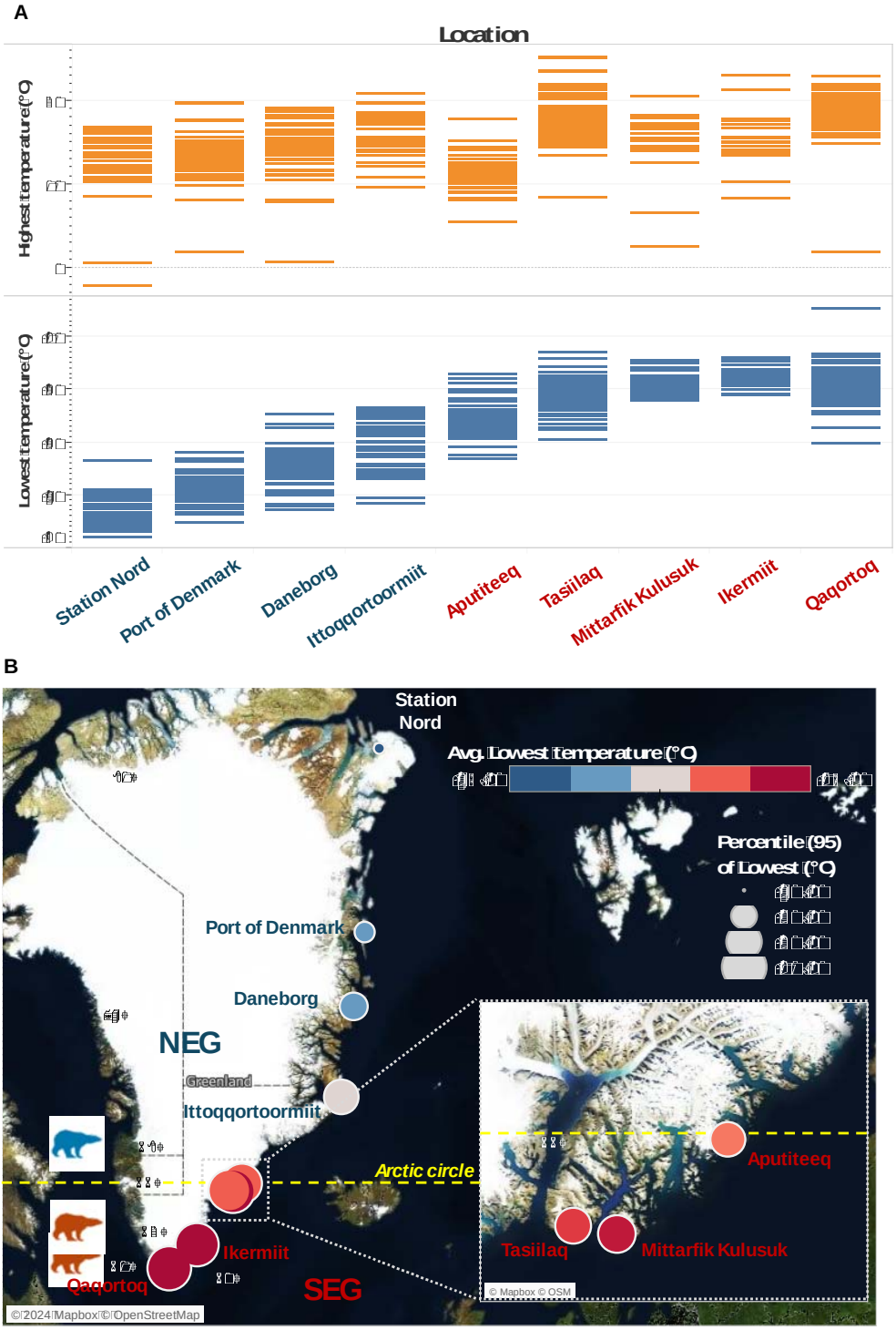


Figure 1- Meteorological observation data of temperature from the DMI. **(A)** Highest and lowest observed temperatures from 1958-2024. **(B)** Colour of points indicates the average lowest observed temperature, with the size of point showing the variance at the 95th percentile. SEG locations included: Aputiteeq, Tasiilaq, Mittarfik Kulusuk, Ikermit and Qaqortoq. NEG locations included: Ittoqqortoormiit, Daneborg, Port of Denmark and Station Nord. In this study north of 64° was considered NEG. The bear icons are indicative of the latitude where they were sampled, red meaning SEG and blue NEG. The Arctic circle is denoted by a dashed yellow line at 66°.

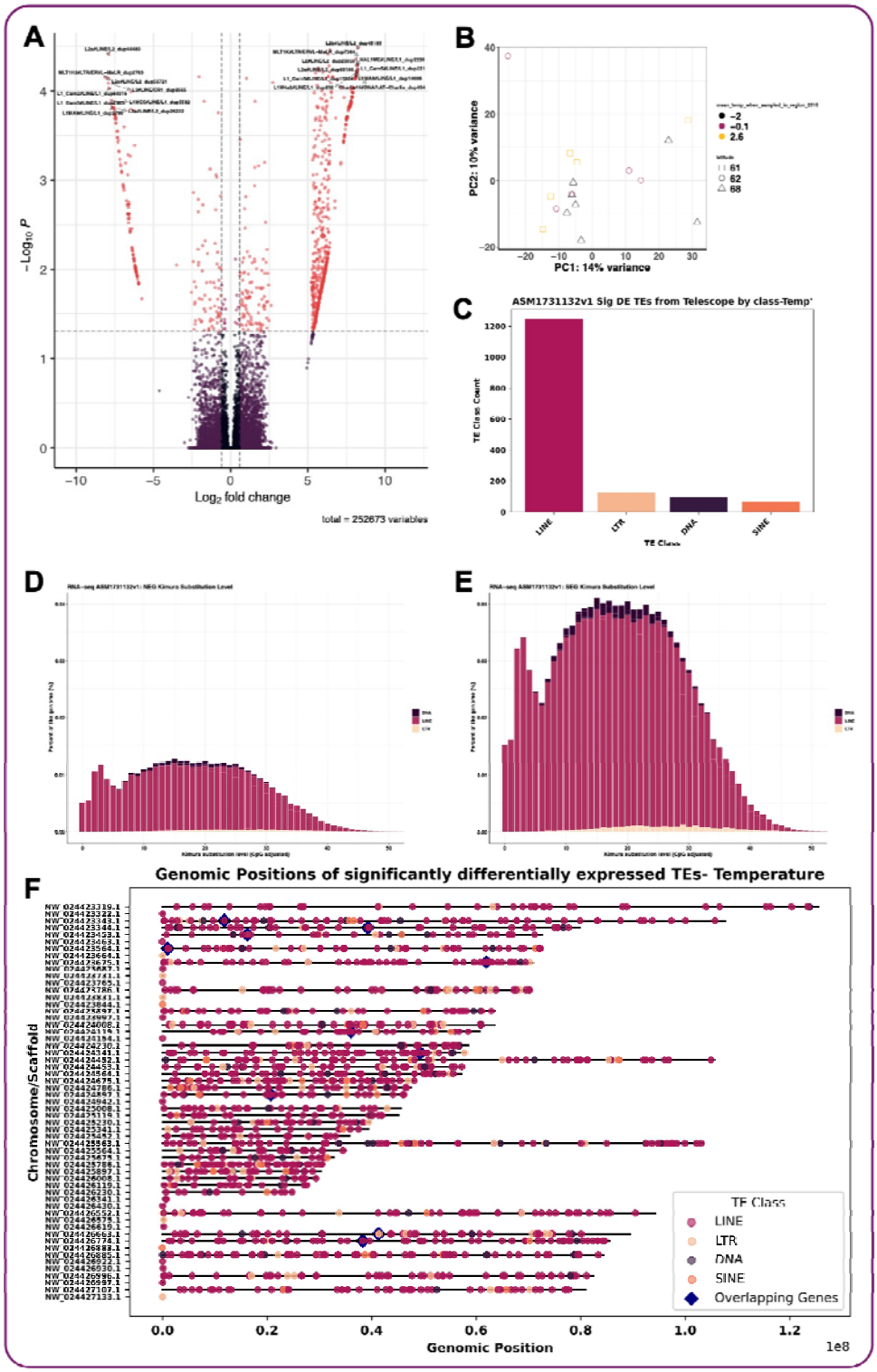


Figure 2- Expression and activity analysis of TEs in SEG v NEG bear to assay the main impact of temperature on TEs. (A) Count data of significantly differentially expressed TEs at the family level. (B) Differential expression analysis of TE species, P = < 0.05 for significance showed 173 significantly differentially expressed TEs. (C) Genomic position of TE species that were significantly differentially expressed. (D-E) TE age and divergence estimation by Kimura substitution analysis

with RepeatMasker. The estimated divergence for SEG bears (E) was greater than that in NEG bears (D).

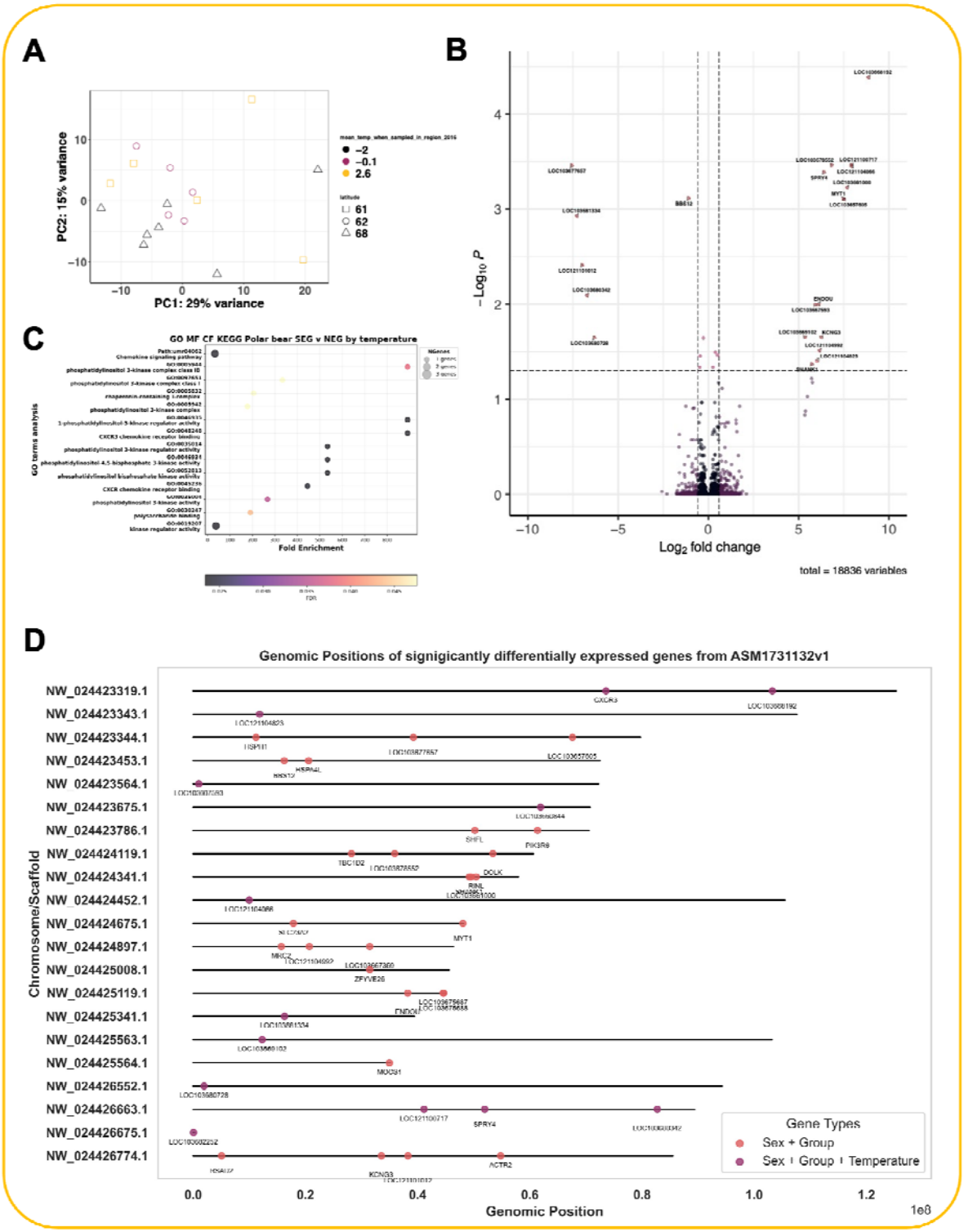
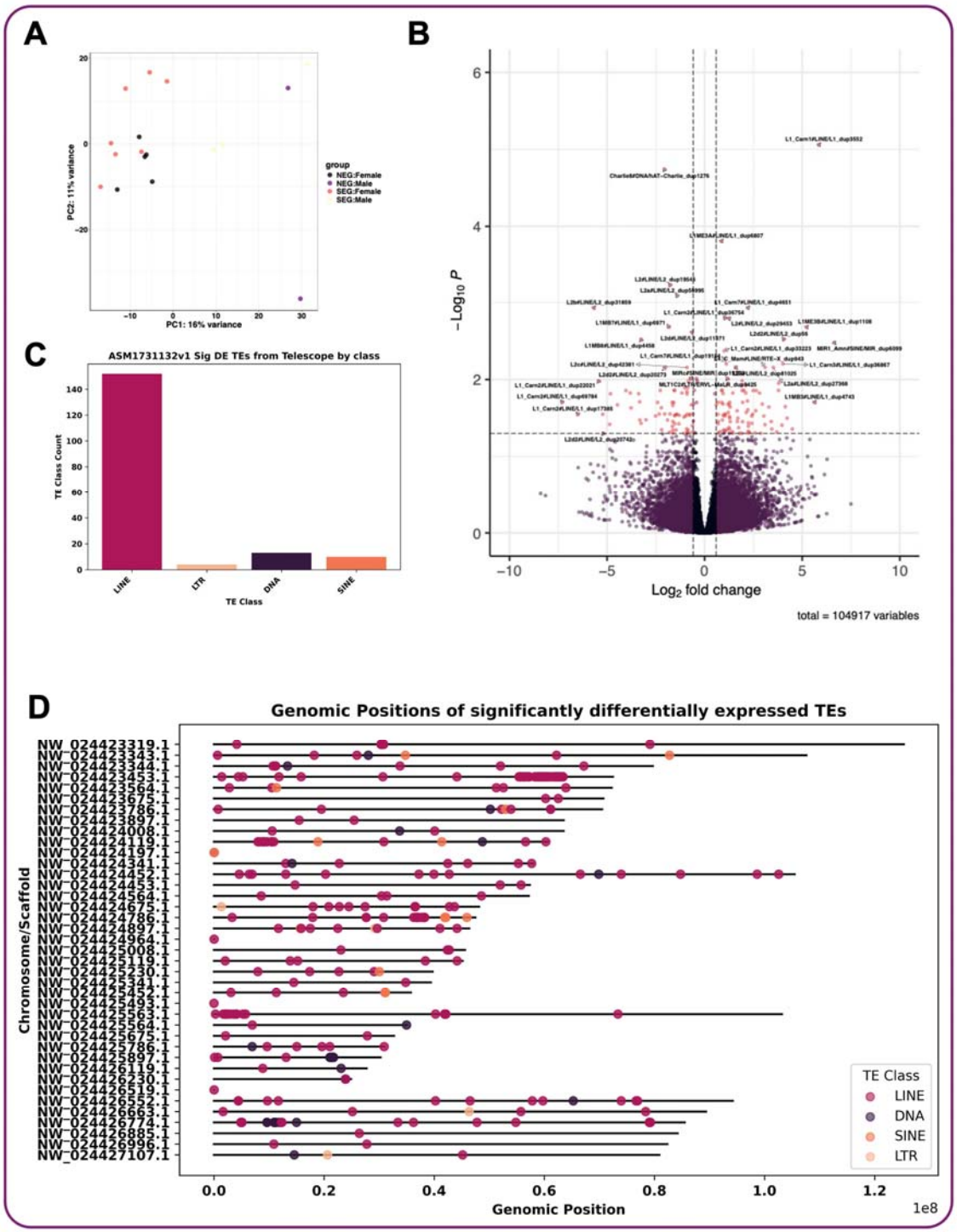


Figure 3 – Analysis of differential gene expression to assay the main impact of temperature on the genomes. (A) PCA analysis of NEG v SEG bears show clustering based on geographical location of sample for gene expression based on the additive model that includes the impact of temperature on

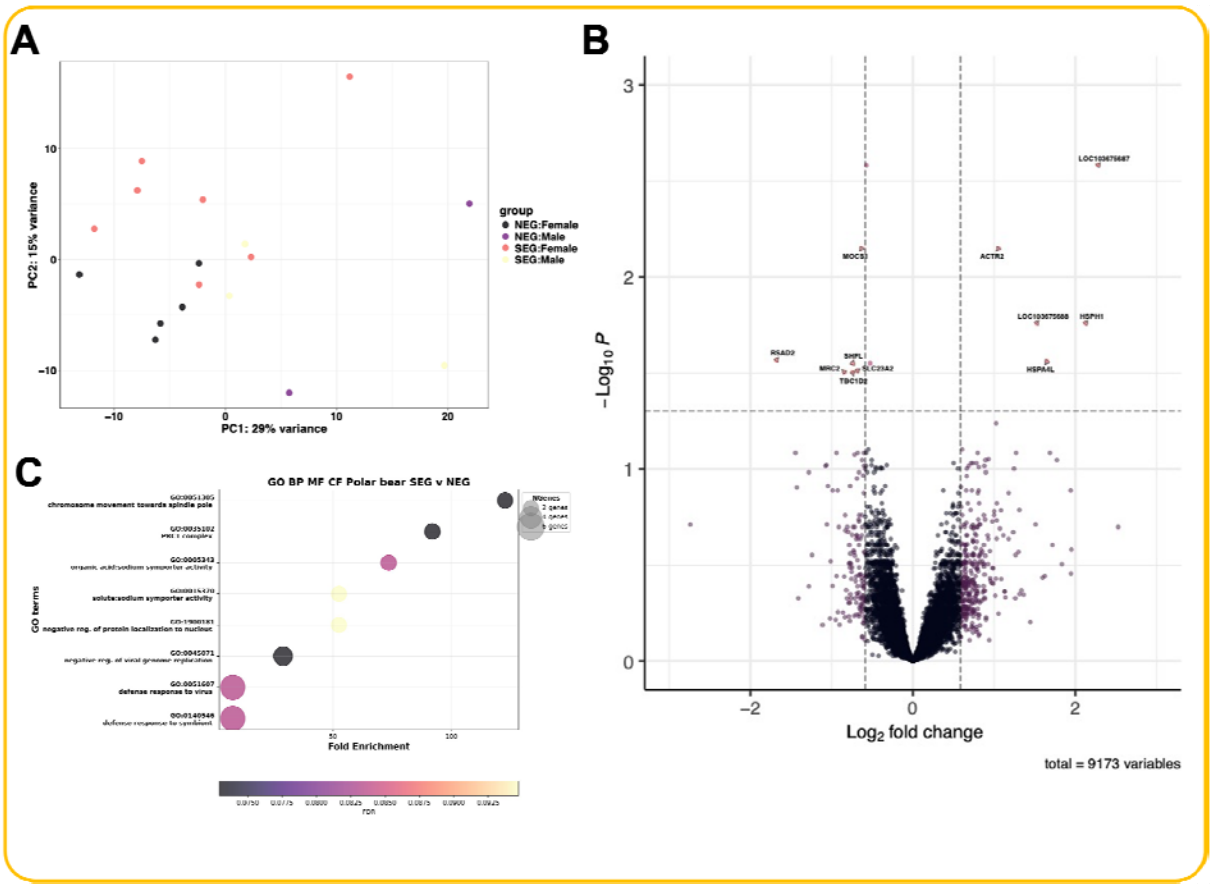
expression. (B) Volcano plot of differentially expressed genes. (C) GO terms analysis of significantly differentially expressed genes. (D) Loci of significantly differentially expressed genes with and without the temperature covariate in DESeq2 analyses.

SUPPLEMENTARY FIGURES



Supplementary Figure 1- Expression and activity analysis of TEs in SEG v NEG bear. (A) PCA analysis of NEG v SEG bears show clustering based on geographical location of sample for TE

expression (B) Differential expression analysis of TE species, P = < 0.05 for significance showed 173 significantly differentially expressed TEs. (C) Count data of significantly differentially expressed TEs at the family level. (D) Genomic position of TE species that were significantly differentially expressed.



Supplementary Figure 2 – RNA-Sequencing data from samples in PRJNA669153 aligned to the reference genome ASM1731132v1 in NEG v SEG bears. (A) Principal component analysis shows variation in gene expression correlates to geographical location of the sample. (B) Volcano plot of the differentially expressed genes observed following DESeq2 analysis, with 13 significantly differentially expressed genes in total (C) GO terms analysis, examining biological processes (BP), Molecular function (MF) and cellular function (CF). All genes expressed in the study were used as a background list, and those with P = < 0.1 processed for GO terms analysis with ShinyGO.

SUPPLEMENTARY FILES

Supplementary File 1- Meteorological data accessed from DMI **–** "Suppl. File. 1- Temperature_data - DMI.csv"

Supplementary File 2- Metadata of samples used for the RNA-seq analysis "Suppl. File. 2. bear_metadata_adult.csv"

Supplementary File 3 – Raw counts RNA-seq ASM1731132v1: "Suppl. File 3-Raw counts RNA-seq ASM1731132v1.tsv"

Supplementary File 4 – DESeq2 data RNA-seq ASM1731132v1: “Suppl. File 4-DESeq2 data RNA-seq ASM1731132v1.csv”

Supplementary File 5 – Telescope raw counts RNA-seq- TEs ASM1731132v1: “Suppl. File 5-Telescope raw counts RNA-seq- TEs ASM1731132v1.csv”

Supplementary File 6 – Telescope DESeq2 data RNA-seq ASM1731132v1: “Suppl. File 6-Telescope DESeq2 data RNA-seq ASM1731132v1.csv”

Supplementary File 7- DESeq2 data RNA-seq ASM1731132v1 Temperature effect: “Suppl. File. 7-Temperature_effect_results_newgen_bear.csv”

Supplementary File 8- DESeq2 data RNA-seq ASM1731132v1 Temperature effect on TE expression: “Supp. File. 8- Temperature_effect_results_newgen_bear_TelescopeTEs.csv”