Supplemental Information for:

Establishment of a non-invasive age estimation method based on fecal DNA methylation levels in brown bears

Satoshi Ohara, Shiori Nakamura, Kyogo Hagino, Yuu Yoshimi, Naoya Matsumoto,
Masami Yamanaka, Masanao Nakanishi, Sei Watanabe, Hideyuki Ito,
Miho Murayama-Inoue, Toshio Tsubota, Jumpei Yamazaki, Michito Shimozuru

Table of Contents:

Supplementary Table S_M1. Sample list for methylation level analysis.

Captive/Wild	Sampling date	Bear ID	Age (years)	Sex	Analysis
Captive	2025/7/3	M0A	0.416	M	C
Captive	2025/7/3	M0B	0.416	M	C
Captive	2025/7/3	Ashiri	1.441	M	C
Captive	2025/7/3	Ashiri	1.441	M	F
Captive	2025/7/3	Ashiri	1.441	M	E
Captive	2025/7/3	Ashiri	1.441	M	E
Captive	2025/7/3	Ehu	2.441	F	F
Captive	2025/7/3	F2B	2.447	F	C
Captive	2025/7/3	F2B	2.447	F	E
Captive	2025/7/3	M2B	2.447	M	C
Captive	2025/7/3	M2B	2.447	M	E
Captive	2025/7/3	F2A	2.452	F	E
Captive	2025/7/3	M2A	2.452	M	C
Captive	2025/7/3	Alen	3.479	M	C
Captive	2025/7/3	Lucky	10.416	F	C
Captive	2025/7/3	Lucky	10.416	F	F
Captive	2025/7/3	Ana	10.477	F	F
Captive	2025/7/3	Sasuke	10.477	M	C
Captive	2025/7/3	Sasuke	10.477	M	E
Captive	2025/7/3	Tsurara	12.397	M	C

Captive	2025/7/3	Maron	12.416	F	C
Captive	2025/7/24	Dai	14.474	M	C
Captive	2025/7/24	Dai	14.474	M	E
Captive	2025/7/24	Dai	14.474	M	E
Captive	2025/7/24	Towa	17.474	F	C
Captive	2025/7/24	Towa	17.474	F	E
Captive	2025/7/3	Sumiri	19.408	F	C
Captive	2025/7/3	Sumiri	19.408	F	F
Captive	2025/7/3	M19A	19.43	M	C
Captive	2025/7/3	M19B	19.43	M	E
Captive	2025/7/3	Becky	23.416	F	C
Captive	2025/7/3	Becky	23.416	F	F
Captive	2025/7/3	Moriko	24.416	F	C
Captive	2025/7/3	Moriko	24.416	F	F
Captive	2025/7/3	Red	26.334	M	C
Captive	2025/7/3	Red	26.334	M	F
Captive	2025/7/3	Suzuri	29.444	F	F
Captive	2025/7/3	Regalia	30.384	F	F
Captive	2025/7/3	Shinko	34.416	F	F
Captive	2025/7/3	Raisa	34.436	F	F
Wild	2023/8/22	23BH01	0.553	F	C
Wild	2023/9/30	23BH02	0.66	F	C
Wild	2024/7/4	23BH02	1.421	F	E

Wild	2022/9/27	21BH08	1.652	F	C
Wild	2022/11/2	21DH01	1.751	M	C
Wild	2022/11/8	21SS02	1.767	M	C
Wild	2023/7/16	21SS03	2.452	F	C
Wild	2023/8/10	ON	2.521	M	E
Wild	2023/8/10	ON	2.521	M	E
Wild	2022/9/27	20CH18	2.652	M	C
Wild	2022/5/5	20CH20	3.255	F	E
Wild	2022/6/3	20SS01	3.334	M	C
Wild	2024/6/4	CY	3.339	F	F
Wild	2022/7/5	LY	3.422	F	E
Wild	2023/8/14	20CH14	3.532	M	F
Wild	2023/8/15	20CH05	3.534	F	F
Wild	2022/9/27	LY	3.652	F	E
Wild	2022/10/19	LY	3.712	F	E
Wild	2022/10/31	20CH20	3.745	F	C
Wild	2023/11/5	20CH14	3.759	M	C
Wild	2022/6/6	BU	4.342	F	F
Wild	2023/7/5	GC	4.422	F	F
Wild	2023/8/23	LY	4.556	F	E
Wild	2023/10/18	LY	4.71	F	C
Wild	2023/7/4	BU	5.419	F	F
Wild	2022/7/10	18SH03	5.436	F	C

Wild	2022/10/19	18SH03	5.712	F	E
Wild	2021/6/2	YN	7.332	F	F
Wild	2021/6/22	BT	7.386	F	F
Wild	2024/10/1	18SH03	7.664	F	E
Wild	2023/7/4	CR	8.419	F	E
Wild	2023/7/6	CR	8.425	F	E
Wild	2025/8/1	18SH03	8.496	F	F
Wild	2023/9/30	CR	8.66	F	C
Wild	2023/6/6	DS	9.342	F	C
Wild	2023/7/5	EI	9.422	F	C
Wild	2023/8/10	EI	9.521	F	F
Wild	2024/5/28	17SS01	10.319	F	C
Wild	2024/10/1	17SS01	10.664	F	E
Wild	2024/11/5	BT	10.759	F	C
Wild	2024/11/5	KP	10.759	F	C
Wild	2022/7/5	GI	12.422	F	C
Wild	2022/7/5	GI	12.422	F	E
Wild	2022/11/8	HC	15.767	F	C
Wild	2024/7/4	HC	17.421	F	E
Wild	2023/9/14	DR	19.617	F	C
Wild	2023/9/14	DR	19.617	F	F
Wild	2021/10/14	RI	19.699	F	C
Wild	2021/10/25	WK	19.729	F	C

Wild	2021/6/15	KR	20.367	F	F
Wild	2022/10/17	WK	20.707	F	E
Wild	2022/11/7	WK	20.764	F	E
Wild	2022/11/7	WK	20.764	F	E
Wild	2023/8/22	KR	22.553	F	C
Wild	2024/7/3	LI	28.418	F	C

Sex: M, male; F, female

Analysis: C, used for model construction; E, used for model evaluation; F, methylation level analysis was failed

Supplementary Table S_M2. Primer information and PCR conditions.

Adjacent gene	Step	Annealing temperature (°C)	PCR product size (bp)	Sequence 5' to 3'
DLX5	1	54	96	F: GGGTGGGGGAATAAAGGTATA
				R: CCACCTAAAATCCCAAAATACCTAC
	2	56	116	F: GGGTGGGGGAATAAAGGTATA
				R: GGGACACCGCTGATCGTTTACCACCTAAAATCCCAAAATACCTAC
				B: GGGACACCGCTGATCGTTTA
	S			GTTATAGGGTTATTTAGAGGT
SLC12A5	1	56	121	F: GGGTGGTTTGGAGTTTTGT
				R: CTACCCCCRAACTATAAACAAT
	2	58	141	F: GGGTGGTTTGGAGTTTTGT
				R: GGGACACCGCTGATCGTTTACTACCCCCRAACTATAAACAAT
				B: GGGACACCGCTGATCGTTTA
	S			AGGAGTTGTTAGAGGTGTTGGA
F2 Intorn12		60	101	F: AGGGTAAGTCCAAGGCATGAC
				R: GCTGGTGGGTAGAGCTTCTGG

F: forward primer; R: reverse primer; B: 5'-biotin-modified primer; S: sequencing primer; Mixed base: R = G+A

Supplementary Table S_M3. Details of the synthetic F2 intron 12 sequence used for host DNA copy number quantification by qPCR.

Gene name	F2_Intron_12
Sequence length (bp)	101
Inserted vector	pUCFa
Target region (Brown bear)	NW_020656153.1, 1,273,581–1,273,681
Target sequence	AGGGTAAGTCCAAGGCATGACAGCAGTACCCTGTCAGGCAACTATTTCTGAAAACAGTCCCTCCTTACTCACCCCCGCTCCCAGAAGCTCTACCCACCAGC

Supplementary Table S_R1. Workflow of the study and changes in the numbers of samples and bears at each stage.

	Wild bears							**Captive bears**						
	Success[a]		**Failure**[b]		**No. Samples**	**Bears**[c]	**Age**[$]	**Success**		**Failure**		**No. Samples**	**Bears**[#]	**Age**[$]
	3/3	**2/3**	**1/3**	**0/3**				**3/3**	**2/3**	**1/3**	**0/3**			
Sampling					1307							112	33	0–34
↓														
Microsatellite					1005	255						72	26	0–34
↓														
Purification					55	31	0–28					40	26	0–34
↓														
First PCR	44	3	5	3	47[e]	27	0–28	28	6	4	2	34[e]	22	0–34
↓														
Second PCR	30	13	3	1	43[f]	26	0–28	26	6	2	0	32[f]	21	0–26
↓														
Pyrosequencing	30	13	0	0	43[g]	26	0–28	14	14	3	1	28[g]	20	0–26
↓														
qPCR (>1392 copies)					37	25	0–28					24	18	0–26
↓														
Age estimation models					25	25	0–28					18	18	0–26
Model performance evaluation					18	10	1–20					10	8	1–19

[a] Success (3/3 and 2/3) indicates number of samples for which all or two among three replicates were successfully analyzed for PCR or pyrosequencing.
[b] Failure (1/3 and 0/3) indicates number of samples for which one or none among three replicates were successfully analyzed for PCR or pyrosequencing. [c] Bears: Number of unique bears. [d]Age: Range of ages of the bears. [e, f, g] Number of successful samples for first-, second-PCR, and pyrosequencing, respectively.

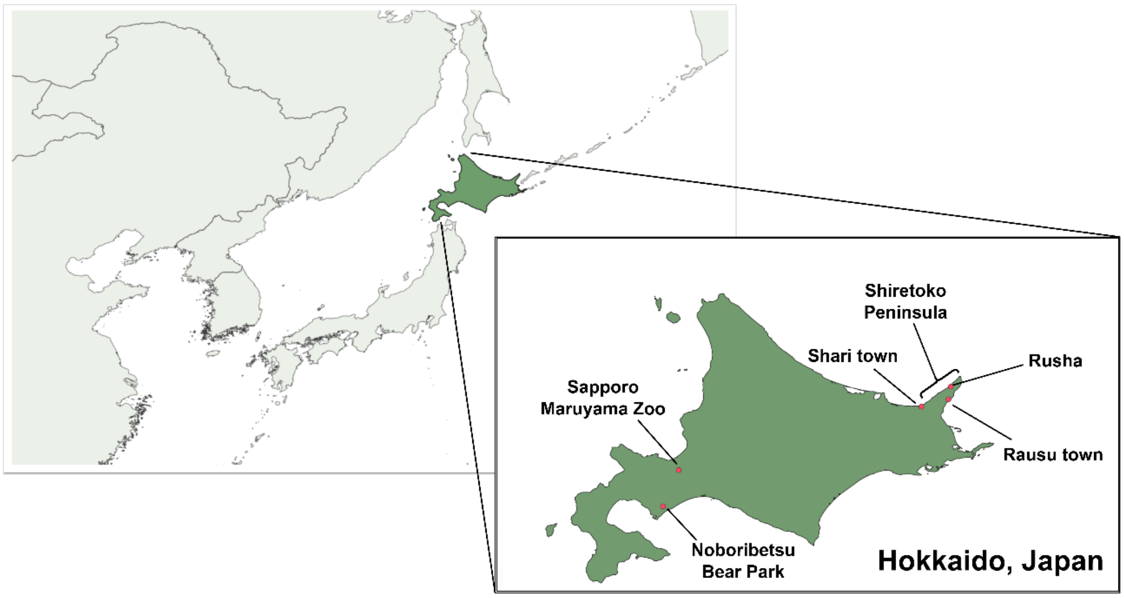
Supplementary Table S_R2. Regression coefficients of the Elastic Net model.

Variable	Coefficient (β)	Interpretation
Intercept	2.32E-11	
DLX5-1	0.1033	Positive correlation with age
DLX5-2	0.1577	Positive correlation
DLX5-3		Not retained in model
SLC12A5-1	0.0990	Positive correlation
SLC12A5-2	0.1414	Positive correlation
SLC12A5-3	0.1628	Positive correlation
SLC12A5-4	0.2348	Positive correlation

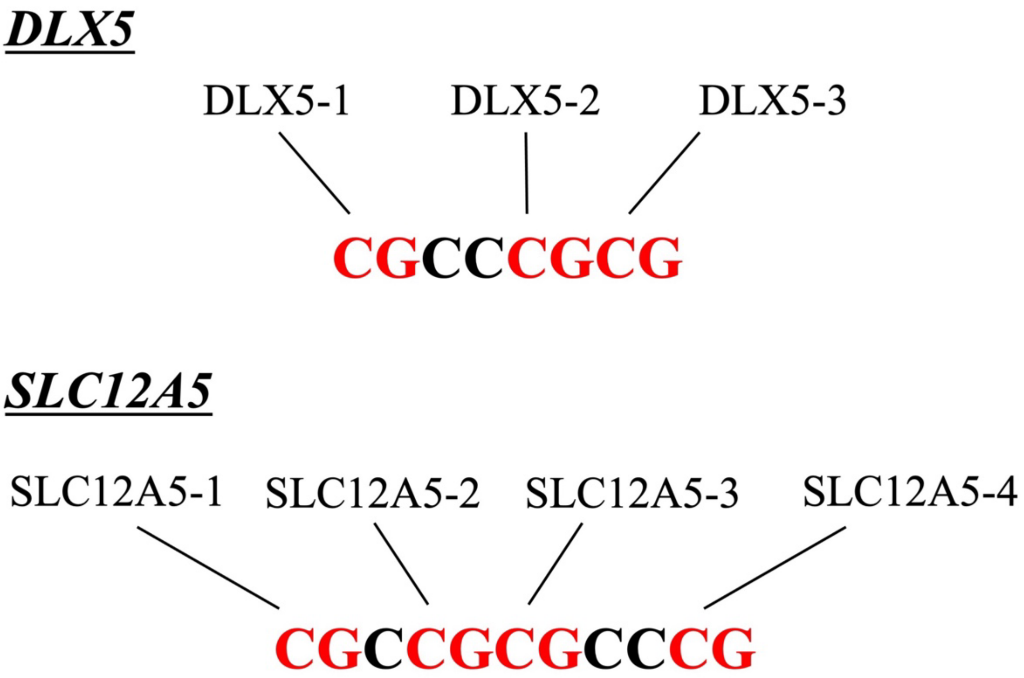
Supplementary Table S_R3. Optimized parameters for the elastic net regression and SVR models.

Model	Used CpGs*	Alpha	Lamda	Cost	Gamma	Epsilon
Elastic net regression	D-1, -2 S-1, -2, -3, -4	0.07	0.1623554			
Support vector regression	D-1, -2, -3, S-1, -2, -3, -4			0.6309573	0.03162278	0.1
	D-1, -2, S-1, -2, -3, -4			0.6309573	0.03162278	0.1
	D-1, -2, S-2, -3, -4			0.6309573	0.03981072	0.1
	D-2, S-2, -3, -4			0.6309573	0.06309573	0.1
	D-2, S-3, -4			0.6309573	0.06310	0.1
	D-1, -2, -3			3.162278	0.1995262	0.1
	D-1, -2			19.95262	0.1995262	0.1
	S-1, -2, -3, -4			79.43282	0.1258925	0.1
	S-2, -3, -4			0.7943282	0.07943282	0.1
	S-3, -4			0.7943282	0.1	0.1

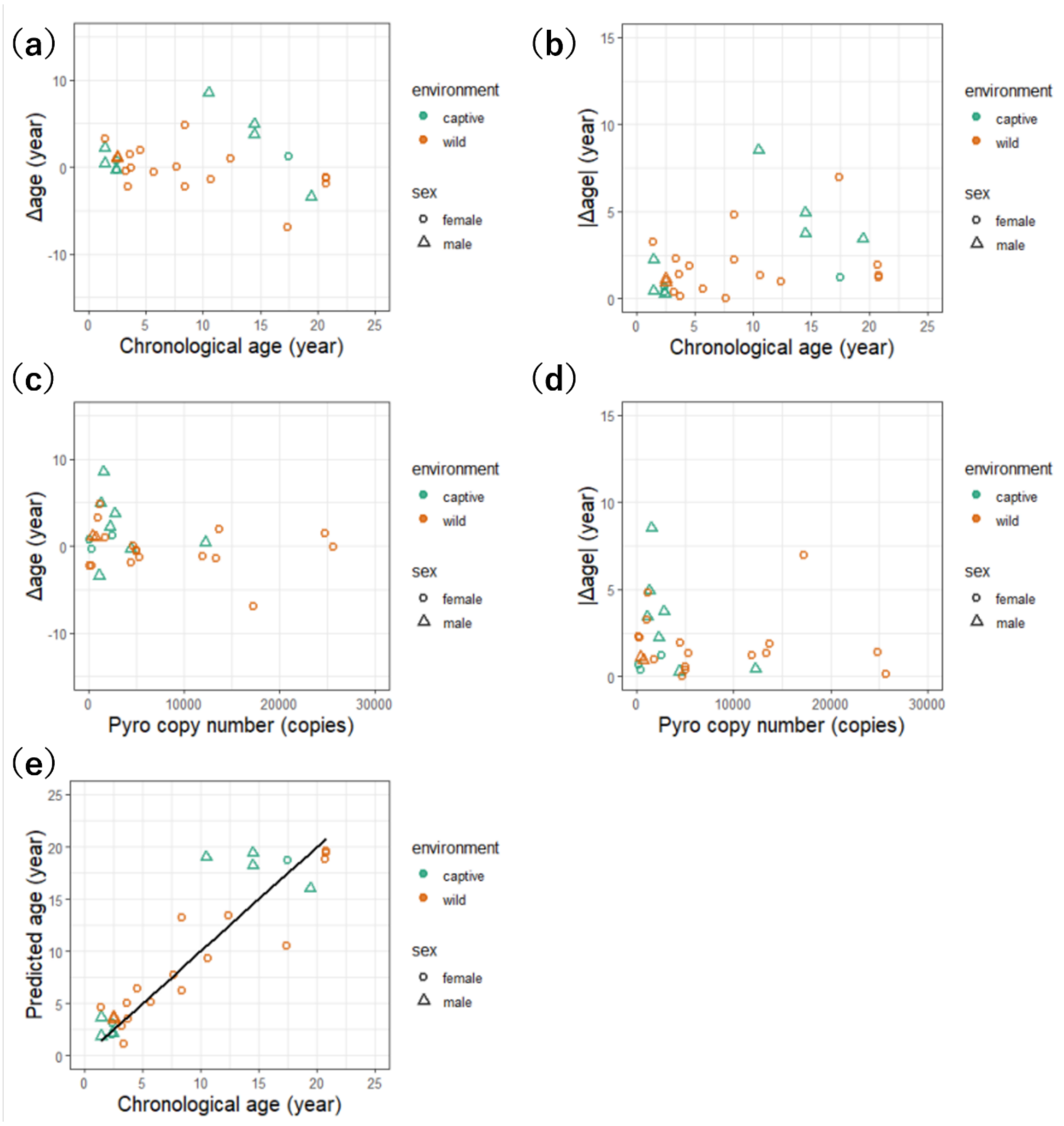
*D and S represent DLX5 and SLC12A5, respectively.



Supplementary Figure S_M1. Map of Hokkaido, Japan, showing Noboribetsu Bear Park, Sapporo Maruyama Zoo and the Rusha area of the Shiretoko Peninsula.



Supplementary Figure S_M2. The nucleotide sequences used for pyrosequencing analysis and the names of the CpG sites used in this study. CpG sites are indicated in red.



Supplementary Figure S_R1. Scatter plot of Δage (years) versus chronological age (years) in captive and wild brown bears (a). Scatter plot of |Δage| (years) versus chronological age (years) in captive and wild brown bears (b). Scatter plot of Δage versus the number of host DNA copies added to the first PCR (c). Scatter plot of |Δage| (years) versus the number of host DNA copies added to the first PCR (d). Scatter plot of predicted age (years) versus chronological age (years) (e). The solid line represents predicted age = chronological age. All of the above data were generated by using samples not included in the model construction.