Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

LINE-1 rewires GPR52 regulation in the rapid evolution of canine tameness

Hua Chen

chen_hua@fudan.edu.cn

Fudan University https://orcid.org/0000-0002-9829-6561

Shilei Zhao

Shanghai Medical College, Fudan University

Yang Yang

China National Center for Bioinformation and Beijing Institute of Genomics, Chinese Academy of Sciences

Yanhu Liu

State Key Laboratory of Genetic Evolution & Animal Models and Yunnan Key Laboratory of Molecular Biology of Domestic Animals, Kunming Institute of Zoology, Chinese Academy of Sciences

Yuedong Zhang

Kunming Institute of Zoology, Chinese Academy of Science

Suiyuan Yang

China National Center for Bioinformation and Beijing Institute of Genomics, Chinese Academy of Sciences

Zhong-Yin Zhou

State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences

Yun Yu

Kunming Institute of Zoology, Chinese Academy of Sciences

Bo-Wen Zhou

State Key Laboratory of Genetic Evolution and Animal Models, Kunming Institute of Zoology, Chinese Academy of Sciences

Lan Gao

Shanghai Medical College, Fudan University

Jin-Xiu Li

State Key Laboratory of Genetic Evolution and Animal Models, Kunming Institute of Zoology, Chinese Academy of Sciences

Yurong Luo

China National Center for Bioinformation and Beijing Institute of Genomics, Chinese Academy of Sciences

Xianchao Ji
Shanghai Medical College, Fudan University
Guo-Dong Wang
Kunming Institute of Zoology, Chinese Academy of Sciences https://orcid.org/0000-0002-9407-4363
Ya-Ping Zhang
Kunming Institute of Zoology https://orcid.org/0000-0002-5401-1114

Article

Keywords: canine domestication, local ancestry, transposable elements, GPR52, neural regulatory evolution

Posted Date: September 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10419682/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

LINE-1 rewires *GPR52* regulation in the rapid evolution of canine tameness

Shilei Zhao[1,2,†], Yang Yang[2,4,†], Yan-Hu Liu[3,5,†], Yue-Dong Zhang[3,5,†], Suiyuan Yang[2,4], Zhong-Yin Zhou[3,5], Yun Yu[3,5,] Bo-Wen Zhou[3,5], Lan Gao[1], Jin-Xiu Li[3,5], Yurong Luo[2,4], Xianchao Ji[1,2], Guo-Dong Wang[3,5], Ya-Ping Zhang[3,5,6*] & Hua Chen[1,2,4*]

[1]Shanghai Medical College, Fudan University, Shanghai, China.
[2]China National Center for Bioinformation and Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing, China.
[3]State Key Laboratory of Genetic Evolution and Animal Models, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China.
[4]School of Future Technology, University of Chinese Academy of Sciences, Beijing, China.
[5]Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, China.
[6]Bio-X Center for Interdisciplinary Innovation, Yunnan University, Kunming, China.

[†]These authors contributed equally: Shilei Zhao, Yang Yang, Yan-Hu Liu and Yue-Dong Zhang.
*Corresponding author. Email: zhangyp@mail.kiz.ac.cn; chen_hua@fudan.edu.cn

Abstract

The rapid behavioral transformation of wolves into dogs, unfolding within an emerging human niche, entailed swift neural adaptation without widespread pleiotropic disruption. To recover adaptive signals embedded in mosaic genomes shaped by complex demographic histories, we developed GAMA, an ancestry-aware framework, and applied it to 1,929 canids. GAMA identified *GPR52* as the top genome-wide selection outlier, enriched for Fertile Crescent–related ancestry. A *Canis*-specific full-length L1-Cf insertion ~39 kb upstream of *GPR52* emerged as the focal candidate cis-regulatory element, associated with ~3.3-fold caudate *GPR52* upregulation, enhancer-like 5′ UTR activity and local chromatin contact. Ancient DNA modeling indicated that L1-Cf persisted as standing variation in wolves, was detected in dogs by ~6.5 ka and rose early in Fertile Crescent–associated populations. Pharmacological GPR52 activation in mice attenuated threat-evoked freezing without broad behavioral impairment. Genome-wide, full-length L1s were enriched near neurodevelopmental genes and contributed to CNS-biased transcription. Together, these findings support standing TE insertions as modular substrates for regulatory exaptation, revealing how mobile elements can reshape neural regulatory programs during rapid behavioral evolution.

Keywords: canine domestication; local ancestry; transposable elements; *GPR52*; neural regulatory evolution.

Main

The transition of gray wolves (*Canis lupus*) into domestic dogs (*Canis lupus familiaris*) represents one of the most rapid and profound behavioral transformations in mammalian history[1,2]. As Charles Darwin first noted, tameness—the heritable attenuation of fear and aggression—constitutes a foundational shift of the domestication syndrome, yet it presents a longstanding challenge to our understanding of behavioral evolution[3,4]. Complex behaviors typically evolve through slow, polygenic shifts to avoid the prohibitive fitness costs of genetic pleiotropy[5,6]. How canine neural circuits were reshaped to favor attenuated fear-reactivity—a prerequisite for survival within an emerging human niche—over an exceptionally compressed timescale remains a central unresolved question at the intersection of evolutionary biology and neuroscience.

Modular regulatory evolution offers a parsimonious route through this constraint, enabling tissue- or circuit-specific adaptation while limiting collateral effects elsewhere[7,8]. Transposable elements (TEs) carry preconfigured regulatory sequences that host genomes can repurpose, providing a reservoir of standing insertions that selection could rapidly exploit as dogs adapted to expanding human-associated niches[9-11]. Yet TE insertions remain underexplored in canine phenotypic evolution.

Furthermore, in dogs, unraveling the genetic basis of phenotypic evolution is compounded by a layered demographic history—serial bottlenecks, recurrent admixture and recent closed-breed formation—that has reshaped allele frequencies and haplotypes, making older adaptive footprints difficult to distinguish from population-history-driven genomic mosaics[12]. To uncover these hidden historical signals, we developed GAMA (Genomic Analysis of Multiple Admixture), an ancestry-aware framework that reconstructs high-resolution local ancestry to identify genomic outliers marked by extreme lineage-specific enrichment.

Applied to 1,929 canid genomes[13], GAMA identified its strongest genome-wide outlier at *GPR52*, a striatally enriched G protein–coupled receptor, driven by enrichment of the Fertile Crescent Guardian–associated ancestry component. Within this locus, a *Canis*-specific full-length L1-Cf insertion upstream of *GPR52* emerged as the focal cis-regulatory element, linking caudate-biased *GPR52* regulation to modular regulatory rewiring during behavioral adaptation. More broadly, this locus points to a wider evolutionary pattern: full-length L1 insertions show non-random population-genetic distributions and CNS-biased transcriptional associations, suggesting that L1-derived regulatory insertions have systematically shaped canine regulatory evolution.

Dropout-adjusted ancient DNA likelihood modeling further revealed a staged amplification-and-dispersal evolution of the L1-Cf locus. The insertion persisted as low-frequency standing variation in wolves but expanded specifically in dogs, appearing in the sampled dog record by ~6.5 ka, with expansion inferred to have begun in Fertile Crescent–Anatolia–Balkans-related populations before wider dispersal. This

trajectory situates the *GPR52*-linked haplotype within the later adaptive history of dogs as early agriculture and expanding human-associated niches reshaped their ecological context[14,15]. Together, our findings identify a standing TE-derived regulatory insertion amplified by selection as a mechanistic bridge from mobile-element variation to striatal gene regulation and attenuated threat reactivity, revealing mobile elements as modular substrates for the rapid remodeling of neural regulatory programs.

Reconstruction of a high-resolution pan-breed ancestry map for 1,929 canids

Modern dog breeds are genomic mosaics assembled from deeply divergent ancestral backgrounds shaped by geographic isolation, regional adaptation and repeated human-mediated dispersal (Fig. 1a). Their complex histories of multi-way admixture, intensified during Victorian breed formation, present a major challenge for local-ancestry inference. To address these challenges, we developed GAMA, an ancestry-aware framework designed for robust local-ancestry inference under complex multi-way admixture and limited reference panels. Simulations showed that GAMA accurately reconstructed ancestry mosaics and consistently outperformed the state-of-the-arts methods including FLARE[16], particularly under extensive multi-way admixture and deeper admixture histories (Supplementary Figs. 1–10).

We applied GAMA to 1,929 canids from the Dog10K dataset, comprising 1,579 modern breed dogs spanning 321 breeds, 281 free-breeding village dogs from 26 populations, 57 gray wolves from 11 populations and 12 mixed-breed individuals[13]. A maximum-likelihood phylogeny resolved gray wolves and 19 representative basal dog lineages, which together defined 20 deeply divergent proxy ancestry components (Supplementary Figs. 11 and 12; Supplementary Table 1). Using this reference panel, we reconstructed a high-resolution pan-breed local-ancestry map, assigning genomic segments across modern breed genomes to these proxy components (Fig. 1b; Supplementary Figs. 13–21).

Notably, the ancestry map revealed a pronounced contribution from a geographic center of early agriculture[17]. The Fertile Crescent Guardian ancestry emerged as the leading component in modern dogs, accounting for 16.66% of pan-breed ancestry (Supplementary Fig. 22). This widespread signature contrasts with localized ancestries such as Kintamani ancestry (0.17%). While the overall gray wolf contribution was limited (1.37%), GAMA recovered high wolf ancestry in breeds with documented recent hybridization, such as the Saarloos Wolfdog (27.22%) and Czechoslovakian Wolfdog (26.98%; Supplementary Fig. 23).

Analysis of 57 wolf genomes further revealed bidirectional but asymmetric dog–wolf ancestry: dog ancestry in wolves (2.99%) exceeded wolf ancestry in modern breeds (1.37%), suggesting an asymmetric pattern of introgression toward wild populations. Furthermore, dog ancestry in wolves exhibited substantial geographic variation—ranging from 5.76% in Iranian wolves near the Fertile Crescent to 1.11% in Swedish wolves (Supplementary Fig. 24)—echoing regional variation in historical opportunities

for human–wolf and dog–wolf contact. Together, these analyses validate GAMA as a basis for identifying ancestry-enriched loci that may have been amplified during human-associated dog evolution.

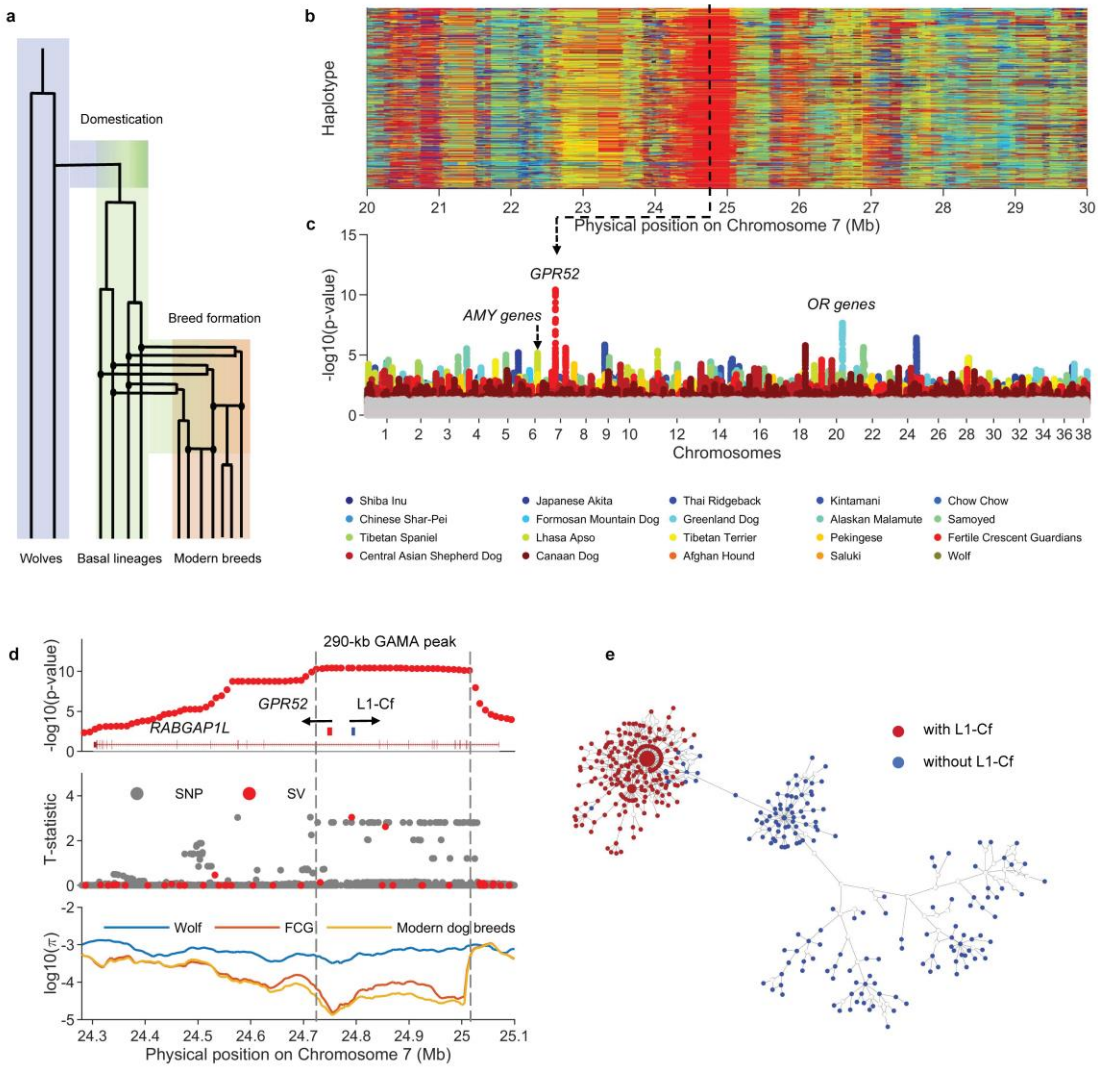


Fig. 1 | Pan-breed local-ancestry reconstruction identifies a *GPR52*-linked ancestry-enriched L1-Cf locus. a, Conceptual framework for using geographically and phylogenetically divergent basal lineages as proxies for ancestral haplotypes. Long-term admixture and recent closed-breed formation reshaped older ancestry components into mosaic modern genomes, whereas basal lineages retain divergent haplotypes that can inform local-ancestry assignment without being treated as direct progenitors. **b,** Local-ancestry map across chromosome 7:20–30 Mb in 1,437 individuals from 284 non-basal modern breeds, excluding mixed/other individuals. Rows represent phased haplotypes and columns represent genomic windows; colors denote inferred proxy ancestry components. The dashed line marks the *GPR52*-linked ancestry-enriched interval. **c,** Genome-wide scan for ancestry-enriched segments across modern breed genomes. Points represent genomic windows and are colored by inferred proxy ancestry component. The 15 components with genome-wide mean ancestry proportion >0.02 are shown as separate color-coded components; windows with local ancestry proportion <0.1 are shown in gray. The strongest signal occurs near *GPR52* on

chromosome 7 and is enriched for the Fertile Crescent Guardian–associated component. **d,** Population-genetic signatures across the *GPR52*-linked ancestry-enriched locus. Tracks show local ancestry enrichment, the branch-specific T statistic for the Fertile Crescent Guardian lineage and nucleotide diversity (π) in wolves, Fertile Crescent Guardians and modern breed dogs. Vertical dashed lines mark the 290-kb GAMA peak interval encompassing *GPR52* and the upstream antisense-oriented full-length L1-Cf insertion. **e,** Haplotype network for the L1-Cf-containing interval upstream of *GPR52* (chr7:24.78–24.80 Mb) in wolves and domestic dogs. Nodes are colored by L1-Cf status, highlighting the clustering of L1-Cf-bearing haplotypes.

Ancestry-specific selection pinpoints a *GPR52*-linked L1-Cf regulatory insertion

The high-resolution GAMA ancestry map identified the strongest genome-wide ancestry-enrichment signal on chromosome 7 (24.72–25.01 Mb). Within this 290-kb interval, 90.90% of haplotypes in modern breeds were assigned to Fertile Crescent Guardian ancestry, far exceeding expectations under the genome-wide null framework (Fig. 1b,c; $P = 3.81 \times 10^{-11}$). To test whether this peak was driven by recent modern-breed introgression within the basal reference lineages, we repeated the scan after masking tracts inferred to be of recent modern-breed origin; this parallel analysis recovered the same chromosome 7 peak interval (Supplementary Fig. 25). Free-breeding village dogs from 26 diverse populations showed a concordant genome-wide significant peak (Supplementary Fig. 26; $P = 7.04 \times 10^{-10}$), with 80.61% of haplotypes assigned to the same ancestry. The presence of this signal in village dogs argues against recent closed-breed formation as the sole driver and suggests that the underlying ancestry-enrichment signal predates modern-breed formation.

The 290-kb candidate interval encompasses two annotated genes, *RABGAP1L* and *GPR52*. *GPR52* emerged as the leading functional candidate because it encodes an evolutionarily conserved orphan G protein–coupled receptor that is highly expressed in the striatum and modulates cAMP-dependent signaling[18,19], providing a plausible route through which selection could have influenced neurobehavioral traits during human-associated dog evolution.

Coding variation was unlikely to explain the ancestry-enrichment signal, as exonic variants within the interval showed only modest allele-frequency differentiation between wolves and dogs. We therefore focused on non-coding variants with predicted regulatory potential. This analysis identified a full-length L1-Cf retrotransposon insertion 38.97 kb upstream of *GPR52* as the most prominent non-coding candidate (Fig. 1d). The insertion showed a pronounced frequency contrast across canids, being rare in wolves but reaching high frequency in Fertile Crescent Guardians and near-fixation in modern breeds, based on population- or breed-level mean frequencies (1.97%, 93.89% and 96.79%, respectively; Supplementary Table 2). Unlike most truncated LINE-1 insertions, this *Canis*-lineage element retains an intact 5′ untranslated region. Because LINE-1 5′ UTRs can harbor regulatory activities capable of influencing proximal gene expression[11,20], L1-Cf represented a strong candidate cis-regulatory element within the ancestry-enriched interval.

Population-genetic analyses further supported selection at the *GPR52* locus. Nucleotide diversity (π) was markedly reduced in Fertile Crescent Guardians and modern breeds relative to gray wolves (Fig. 1d), consistent with ancestry-associated haplotype expansion. Branch-specific T-statistic analysis identified L1-Cf as the most strongly differentiated marker distinguishing the Fertile Crescent Guardian lineage from other basal dog lineages[21]. A haplotype network further revealed a densely clustered topology of L1-Cf-bearing chromosomes, consistent with rapid expansion of the insertion-bearing haplotype (Fig. 1e)[22,23].

We next used 84 Chinese village dogs to further evaluate haplotype-based evidence for selection at the L1-Cf locus. This free-breeding population panel retained the L1-Cf insertion at intermediate frequency (48.21%), providing sufficient variation for allele-specific haplotype analyses. Both Extended Haplotype Homozygosity (EHH) and HaploSweep analyses revealed significantly extended homozygosity tracts associated with the L1-Cf allele (Supplementary Fig. 27)[24,25]. Using HMM-Sweep[26], we estimated a selection coefficient of s = 0.0117 (95% CI, 0.0089–0.0312), providing additional haplotype-based support for positive selection at this locus outside the context of closed modern breeds.

Together, these analyses identify the *GPR52*-linked interval as a candidate signature of selection associated with Fertile Crescent Guardian ancestry and nominate the upstream *Canis*-specific full-length L1-Cf insertion as the focal regulatory variant. This insertion provided a defined molecular entry point for testing how selection reshaped striatal *GPR52* regulation during human-associated dog evolution.

The L1-Cf insertion links caudate-specific *GPR52* upregulation to enhancer-like activity

To determine whether L1-Cf status was associated with altered *GPR52* expression in neural tissues, we analyzed *GPR52* expression across 20 central nervous system (CNS) tissues from the DoGA dataset[27]. After excluding three tissues with fewer than two samples in either comparison group, 17 CNS tissues remained testable. In the caudate nucleus, a major striatal region, *GPR52* expression was markedly higher in homozygous L1-Cf-bearing dog samples than in wolf comparison samples (~3.3-fold; adjusted P = 0.0145; Fig. 2a). This difference was not observed in the amygdala, frontal cortex or other testable CNS regions (Fig. 2a and Supplementary Fig. 28). By contrast, *RABGAP1L*, the flanking gene in the same interval, showed no comparable expression difference across these tissues (Supplementary Fig. 29). These expression patterns support an L1-Cf–*GPR52* regulatory association that is most evident in the caudate nucleus, rather than a broadly shared transcriptional shift across CNS tissues.

We next tested whether the L1-Cf 5′ UTR could provide regulatory activity. Because L1 5′ UTRs can harbor cryptic promoter or enhancer activity[20], we cloned the L1-Cf 5′ UTR and tested it in reporter assays in HEK293T cells, using upstream-of-promoter

and downstream-of-reporter placements in both orientations. The L1-Cf 5′ UTR increased reporter expression in all four positional and orientation configurations (Fig. 2b), a pattern consistent with enhancer-like regulatory potential in a heterologous cellular context.

Finally, we examined the chromatin context surrounding the L1-Cf insertion site in the canine caudate nucleus. Juicer-derived loop calls at 5-, 10- and 25-kb Hi-C resolutions identified a significant local interaction between a bin overlapping *GPR52* and a nearby reference-mappable bin flanking the L1-Cf insertion site, with support across HiCCUPS local background models (FDR $< 1 \times 10^{-15}$; Fig. 2c). Because L1-Cf is repetitive, Hi-C contacts arising from the repeat body itself are difficult to resolve uniquely. The detectable flanking-bin interaction therefore supports a model in which the L1-Cf-containing region participates in the local three-dimensional chromatin context of *GPR52*. Together, the caudate expression association, enhancer-like reporter activity and caudate chromatin-context evidence identify L1-Cf as a focal cis-regulatory candidate linked to caudate-specific *GPR52* upregulation.

GPR52 activation attenuates acute threat-evoked defensive responses in mice

Human GWAS have linked variation near *GPR52* to multiple neuropsychiatric phenotypes[28-30], and rodent studies place GPR52 in striatal dopaminergic circuitry, where its activation elevates cAMP in DRD2-expressing indirect-pathway neurons implicated in aversive learning and avoidance[19,31,32]. As exploratory breed-level behavioral context, we analyzed standardized breed-aggregated traits across 170 dog breeds stratified by L1-Cf frequency (Supplementary Table 3). Because L1-Cf is near fixation in modern breeds, breeds were grouped into a high-frequency group (HFG; n = 159, fixed) and a low-frequency group (LFG; n = 11, frequency $\leqslant$ 0.5). Despite this unbalanced distribution, Bonferroni-corrected comparisons identified higher scores in HFG breeds for three human-rated traits: “best family dog” ($P = 1.48 \times 10^{-7}$), “kid-friendly” ($P = 6.00 \times 10^{-8}$) and “intelligence” ($P = 9.65 \times 10^{-5}$), together with a higher composite “all-around-friendliness” score ($P = 9.16 \times 10^{-6}$).

Given the inherent limitations of breed-level genotype–phenotype associations, we next tested the behavioral consequences of enhanced GPR52 signaling pharmacologically in mice using HTL0041178, a highly selective *GPR52* agonist[33,34]. Because L1-Cf was associated with increased *GPR52* expression and enhancer-like activity in dogs, this approach provided a tractable test of whether increased GPR52 signaling alters acute defensive behavior. We focused on looming-evoked freezing, an assay of visually evoked innate threat reactivity[35,36].

In the looming stimulus test, which mimics an approaching aerial predator, vehicle-treated control mice exhibited immediate and prolonged freezing, a stereotypical innate defensive response. By contrast, GPR52 agonist-treated mice showed significantly shorter freezing durations during the initial exposure (n = 20 per group; $P = 3.14 \times 10^{-3}$; Fig. 2d). This effect persisted across the early trial block (trials 2–5; $P = 4.56 \times 10^{-2}$)

before converging with controls, consistent with habituation. The reduction in freezing was not accompanied by detectable changes in anxiety-like behavior or locomotor/exploratory measures in the elevated plus maze (Supplementary Fig. 31), sociability in the three-chamber sociability assay (Supplementary Fig. 32), or sensorimotor gating in the prepulse inhibition assay (Supplementary Fig. 33). These complementary behavioral assessments argue against broad sedation, locomotor impairment, generalized anxiolysis or disruption of acoustic startle gating as primary explanations for the looming-response phenotype.

Together, these behavioral experiments show that potentiating GPR52 signaling attenuates acute visually evoked defensive freezing without detectable broad effects on anxiety-like behavior, locomotor/exploratory activity, social behavior or sensorimotor gating. This behavioral profile is consistent with a key dimension of tameness: reduced acute threat reactivity without broad impairment of social or sensorimotor functions.

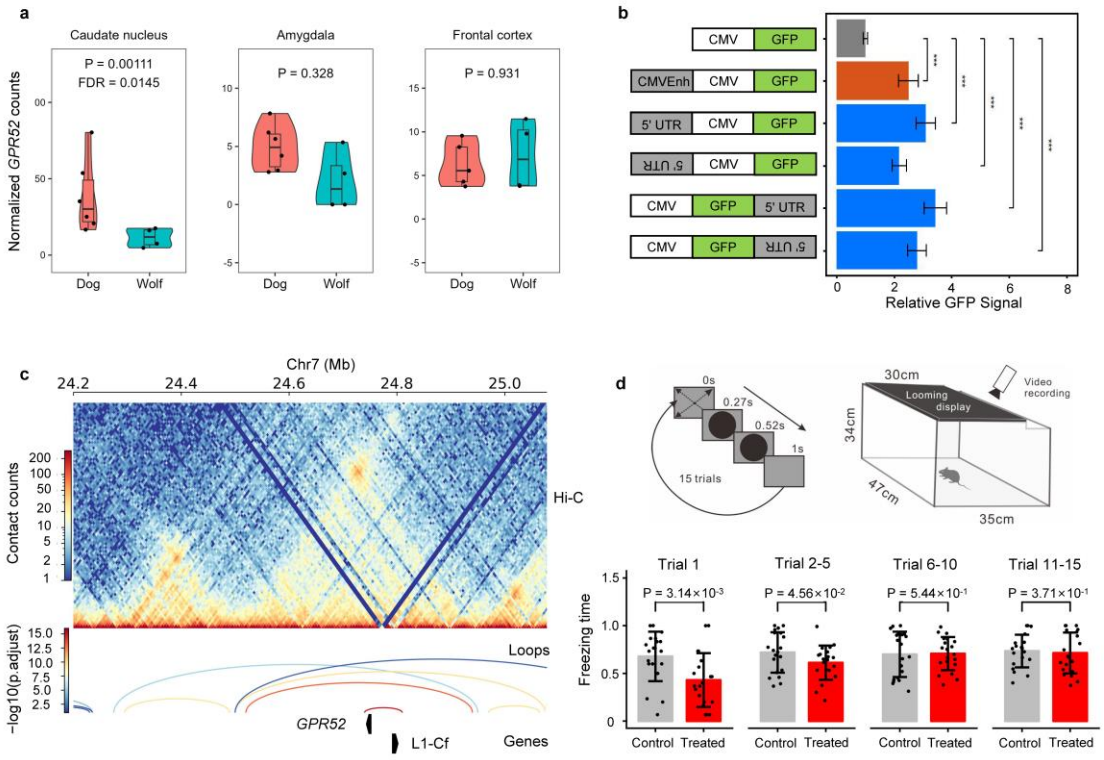


Fig. 2 | L1-Cf is associated with caudate *GPR52* upregulation, enhancer-like activity and reduced defensive freezing. a, *GPR52* expression in selected CNS regions comparing homozygous L1-Cf-bearing dog samples with wolf comparison samples. Wolf samples included confirmed non-carriers and wolves without genome-based L1-Cf calls that were treated as presumed non-carriers given the low L1-Cf frequency in wolves. Normalized counts are shown for caudate nucleus, amygdala and frontal cortex; points represent individuals. Of 20 DoGA CNS tissues, 17 met the testing criterion ($n \geq 2$ per group); only *GPR52* in the caudate nucleus remained significant after DESeq2 Wald testing and FDR correction[37]. Full *GPR52* and *RABGAP1L* comparisons are shown in Supplementary Figs. 28 and 29. **b,** CMV–GFP reporter assays of L1-Cf 5′ UTR activity in HEK293T cells at 48 h. Upstream/downstream and forward/reverse constructs were compared with

promoter-only and enhancer-positive controls; GFP signal was normalized to the promoter-only control. Data are mean ± s.d.; *P < 0.05, **P < 0.01 and ***P < 0.001 by two-tailed Student's t-test. Time-course data are shown in Supplementary Fig. 30. **c,** Caudate Hi-C analysis at the *GPR52* locus. Merged 5-kb, 10-kb and 25-kb Juicer loop calls identified a significant interaction between a *GPR52*-overlapping mappable bin and a mappable bin flanking the L1-Cf insertion site (observed contacts = 94; FDR < 1×10^{-15} across HiCCUPS local background models)[38,39]. **d,** Mouse looming test after vehicle or selective GPR52 agonist HTL0041178 treatment (n = 20 per group). One-sided tests showed reduced freezing during trial 1 and trials 2–5, but not during later trial blocks.

Ancient DNA analyses reconstruct the rise and dispersal of L1-Cf

To distinguish a genuine historical adaptive rise of L1-Cf in dogs from patterns shaped by demographic history, we analyzed a broad spatiotemporal panel of published ancient dog and wolf genomes from Eurasia and the North American Arctic, retaining samples with non-zero local coverage for model-based frequency estimation (Supplementary Table 4). Sample-level L1-Cf detection was based on high-quality 5′ junction reads spanning the boundary between the upstream genomic flank and the L1-Cf insertion[40]; junction-read evidence for all ancient canids classified as positive is shown in Supplementary Figs. 34–36. Because low-coverage ancient genomes are vulnerable to allelic dropout, the absence of junction reads cannot be interpreted as definitive absence. We therefore developed a dropout-adjusted likelihood framework that incorporates local coverage in the 1-kb region upstream of the L1-Cf 5′ UTR, read-length distributions, breakpoint-anchor requirements and genotype-specific detection probabilities. We used this framework for two complementary analyses: to estimate L1-Cf frequencies within temporal bins and to fit a selection-onset trajectory model for inferring onset time and selection intensity. Simulation benchmarks for both modeling components are shown in Supplementary Figs. 37–39.

Of 98 age-dated ancient dog genomes, 73 retained non-zero local coverage upstream of the L1-Cf 5′ UTR and were retained for subsequent analyses. Direct junction-read evidence was not observed in the oldest dog samples, but detection emerged around 6.5 ka BP and became more frequent in later time bins, with early carriers identified in Greece (6,544 BP), Iran (5,826 BP) and Croatia (4,900 BP), within and around the Fertile Crescent–Anatolia–Balkans corridor (Fig. 3a,b). Because non-detection in ancient genomes can reflect low local coverage, short reads or limited breakpoint mappability rather than true absence, we applied the dropout-adjusted bin-level frequency estimator. This analysis inferred an increase in L1-Cf frequency from near zero in the oldest time bin to high levels in more recent ancient dogs (Fig. 3c).

Wolves showed a contrasting pattern. Among 70 age-dated ancient wolf genomes with non-zero local coverage, five showed L1-Cf junction-read evidence, including one from Germany dated to 57,233 BP (Supplementary Fig. 36). Yet the insertion remained rare: of 57 modern wolves, only four carried L1-Cf, each on a single haplotype. Neighbor-joining and local-ancestry analyses further showed that three of these four haplotypes were associated with dog-derived local ancestry (Supplementary Figs. 40 and 41). Thus,

the wolf data support an ancient origin of L1-Cf in canids but provide little evidence for its sustained expansion in wolves.

We next reconstructed where and when L1-Cf frequency first began to rise in ancient dogs. Ancient dogs were grouped into three broad regions: R1, Fertile Crescent–Anatolia–Balkans-related dogs; R2, Western Eurasian dogs; and R3, Eastern/Northern Eurasian and Arctic North American dogs (Fig. 3a; Supplementary Table 4). A dropout-adjusted selection-onset trajectory model fitted separately by region inferred the earliest onset in R1 (8.29 ka BP; 95% CI, 7.54–15.64 ka BP), with a positive selection coefficient (s = 0.019; 95% CI, 0.00485–0.0296), followed by later increases in R2 and R3 (Fig. 3d). Together with the GAMA ancestry-enrichment peak and population-genetic analyses of contemporary genomes, these trajectories confirm selection-driven amplification of the L1-Cf haplotype, placing its earliest detectable rise in R1 within expanding agriculture-shaped niches before wider dispersal.

To test whether this ordering depended on regional definitions or model constraints, we performed two complementary robustness analyses. First, we repeated the region-specific trajectory analysis under an alternative regional definition (R1′–R3′), in which Balkan samples were reassigned from R1 to R2 and Arctic samples were excluded from R3; this preserved an earlier rise in R1′ followed by later increases in R2′ and R3′ (Fig. 3d). Second, we jointly fitted a shared-selection model across the three regions to assess temporal ordering, allowing only initial frequency and onset time to vary by region. Across bootstrap replicates, 93.4% supported the onset-time ordering R1 ≥ R2 ≥ R3 (Supplementary Fig. 42). These robustness tests show that the R1-first ordering is stable to alternative regional definitions and to constraining selection intensity across regions.

Together, these ancient DNA analyses place the modern ancestry-enrichment signal at *GPR52* within an explicit historical trajectory. Rather than a dog-specific de novo event, L1-Cf represents a deeply rooted canid insertion that persisted at low frequency in wolves. In dogs, its temporal rise in ancient genomes, association with Fertile Crescent-associated local ancestry, and R1-first expansion pattern converge on a staged amplification-and-dispersal model. Thus, the key evolutionary signal was not the origin of the insertion itself, but selective amplification of standing TE-derived variation as agriculture-shaped human niches expanded before broader dispersal.

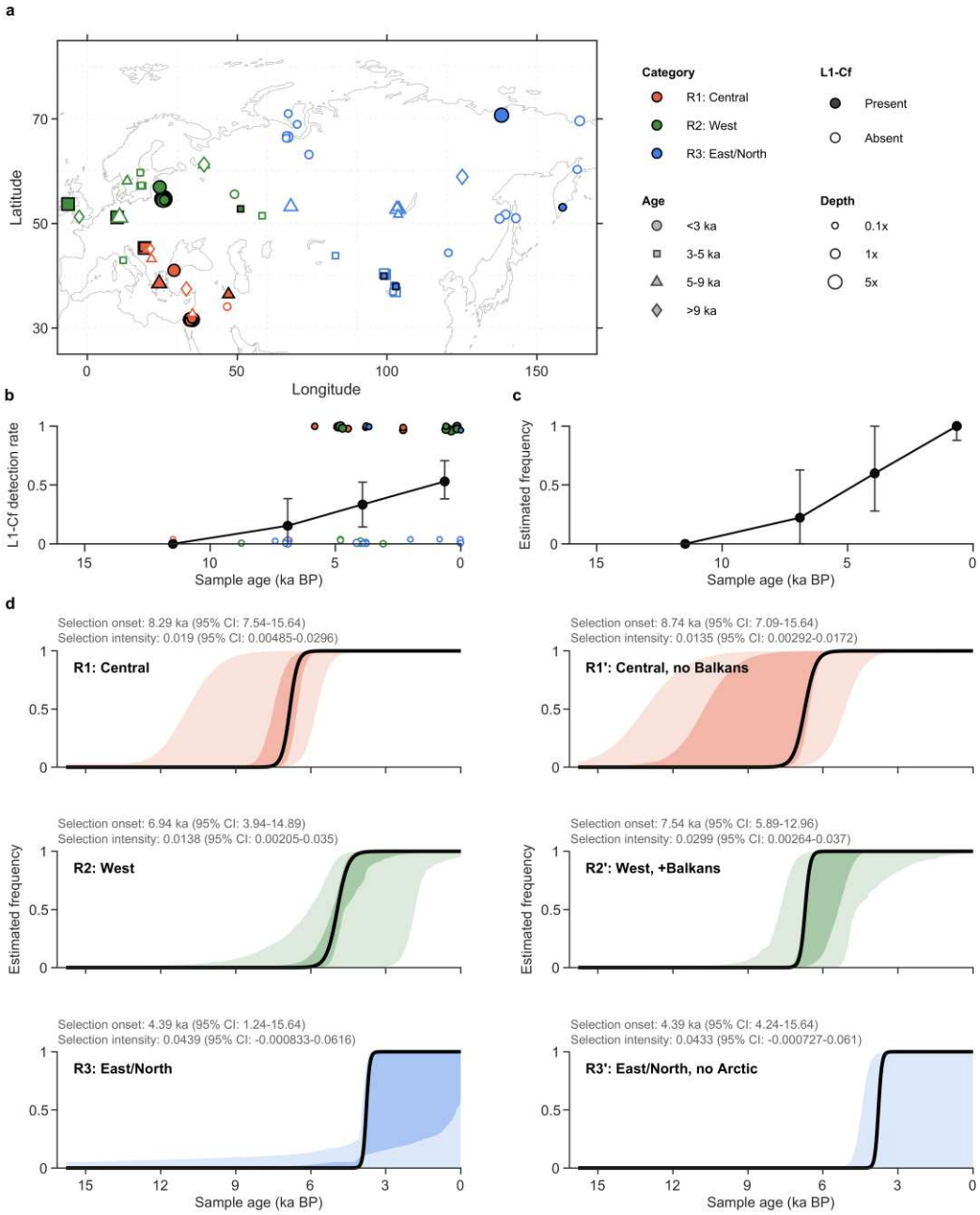


Fig. 3 | Spatiotemporal dynamics of the L1-Cf insertion in ancient dogs. a, Geographic distribution of ancient dog samples included in L1-Cf detection analysis within the Eurasian plotting window. Colors denote regional assignments: R1, Fertile Crescent–Anatolia–Balkans-related dogs; R2, Western Eurasian dogs; and R3, Eastern/Northern Eurasian and Arctic North American dogs. Marker shape indicates age bin, filled and open symbols denote junction-read detection and non-detection, and point size indicates local sequencing depth in the 1-kb region upstream of the L1-Cf 5′ UTR. Samples with zero local depth were excluded; two samples outside the plotting window are not shown but were retained in temporal and model-based analyses. **b,** Sample-level L1-Cf detection across time. Colored points represent individual ancient samples; black points and error bars show bin-level observed detection rates with bootstrap 95% confidence intervals. **c,** Dropout-adjusted L1-Cf frequency estimates across the same temporal bins, accounting for false-negative detection due to local depth, read length and breakpoint mappability. Points show full-data estimates; error bars

indicate bootstrap 95% confidence intervals. **d,** Region-specific frequency trajectories inferred using the dropout-adjusted selection-onset model. Left, primary regional definitions; right, alternative regional definition R1′–R3′, with Balkan samples reassigned from R1 to R2 and Arctic samples excluded from R3. Pale and dark ribbons indicate bootstrap 95% and 50% trajectory intervals; black curves show full-data fits. Text gives full-data estimates and bootstrap 95% confidence intervals for onset time and selection intensity. Ages are in ka BP, with older samples to the left.

Genome-wide full-length L1 analysis reveals adaptive signatures, neural-gene enrichment and CNS-biased regulation

Having established L1-Cf as a focal regulatory candidate at *GPR52*, we asked whether this locus reflected a broader class of L1-associated evolutionary and transcriptional variation in dogs. Across a global *Canis* cohort, we characterized genome-wide L1 insertion patterns and compared the allele-frequency spectrum of full-length L1 insertions (>6 kb) with that of putatively neutral truncated L1 elements. Full-length L1 insertions showed a significant excess of high-frequency alleles relative to truncated elements (χ^2 test, $P = 3.6 \times 10^{-3}$; Fig. 4a), consistent with positive selection acting on a subset of full-length L1 insertions. In this genome-wide context, the *GPR52*-linked L1-Cf insertion showed the strongest wolf–dog frequency shift (Supplementary Fig. 43), placing it as a leading locus-specific candidate in the domestic dog lineage.

We next asked whether L1 insertion neighborhoods were functionally structured. Using a diverse canid pan-genome, we identified 1,213 named genes within 50 kb of L1 insertions for enrichment analysis (Supplementary Tables 5 and 6). These genes were enriched for pathways related to neurodevelopment, synaptic function and neurobehavioral regulation, including G protein-coupled receptor signaling, neuroactive ligand–receptor interaction and synaptic assembly (Fig. 4b; Supplementary Tables 7–9). Of the top 50 full-length L1 insertions exhibiting the largest frequency increases in dogs relative to wolves, several mapped near genes involved in neural-circuit organization (Supplementary Fig. 43). These included *SLITRK* family members (*SLITRK1*, *SLITRK4* and *SLITRK5*), which contribute to synapse formation[41,42], and *EPHA7*, a regulator of dendritic branching[43]. Notably, this outlier set also featured an L1 insertion near *SLC5A1*, consistent with the overlap between dog-enriched L1-associated regions and loci previously implicated in canine dietary adaptation[14].

These functional enrichments were paralleled by CNS-biased transcriptional signatures. Using the DoGA dataset[27], we quantified L1-associated expression shifts at both tissue and organ-system levels, retaining 52 tissues from 11 organ systems for tissue-level analysis and 12 organ systems for system-level analysis. Seven tissues showed significant L1-associated expression shifts (FDR < 0.05), six of which were CNS tissues, including the frontal cortex ($FDR = 9.3 \times 10^{-3}$) and caudate nucleus ($FDR = 2.85 \times 10^{-2}$; Fig. 4c,d; Supplementary Table 10). This CNS predominance was reinforced at the organ-system level, where CNS showed the strongest signal ($FDR = 3.94 \times 10^{-19}$) and a larger proportion of genes with elevated fold changes than other

systems (Fig. 4e; Supplementary Table 11). Moreover, the mean CNS expression shift was significantly greater for genes near dog-dominant than wolf-dominant L1 insertions ($P = 3.44 \times 10^{-5}$; Supplementary Table 12), linking population-level L1 frequency divergence to CNS-biased transcriptional variation.

Together, these genome-wide analyses reveal that the L1-Cf insertion at *GPR52* is representative of a broader class of full-length canine L1 elements exhibiting non-random evolutionary, functional, and transcriptional signatures. Critically, the pronounced CNS bias of these signatures supports a model in which L1 insertions provide a recurrent substrate for regulatory diversification underlying the rapid evolution of tameness and behavioral traits in the domestic dog lineage.

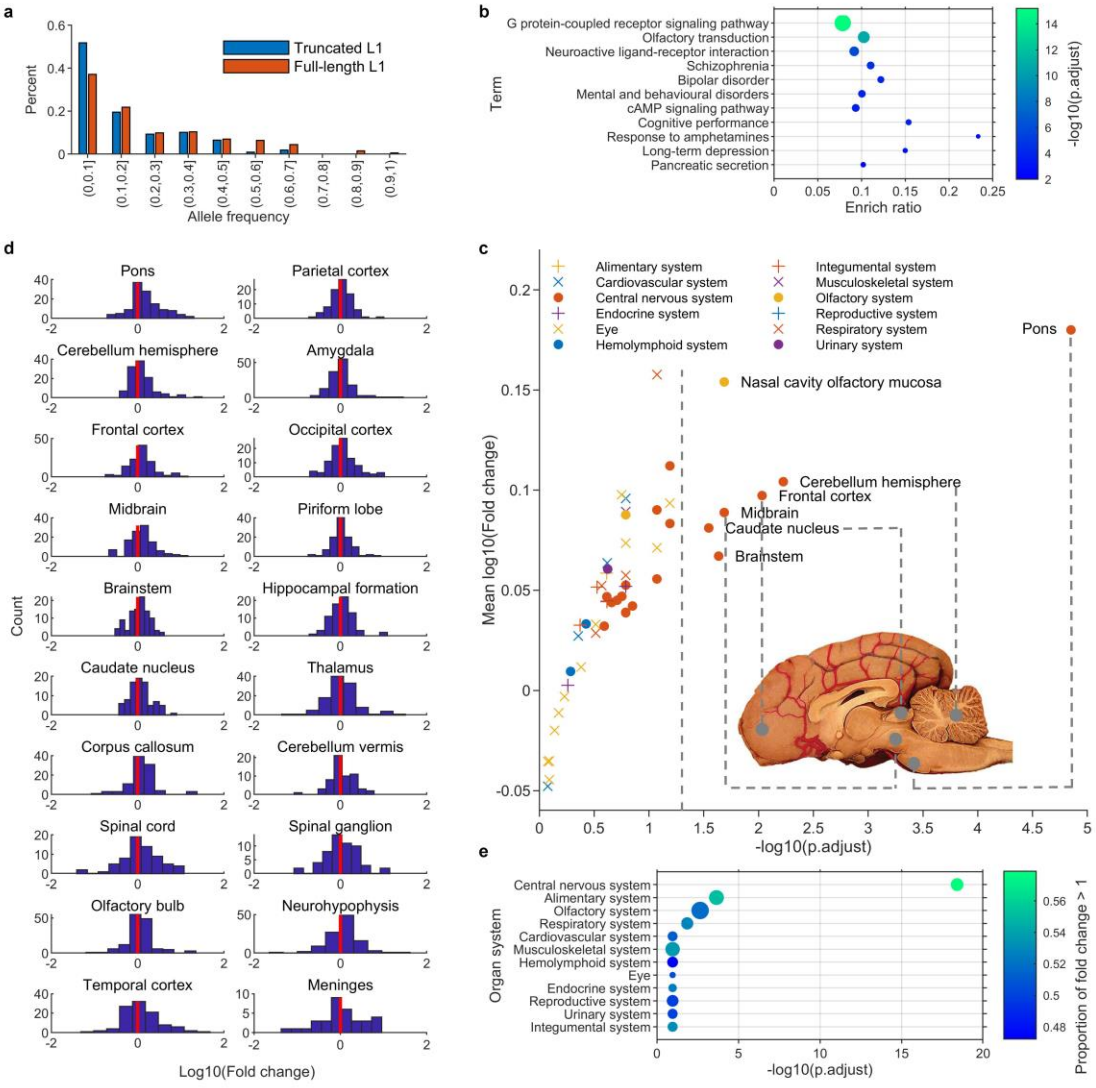


Fig. 4 | Full-length canine L1 insertions show adaptive signatures, neural-gene enrichment and CNS-biased regulatory associations. a, Allele-frequency spectra of full-length L1 insertions (>6 kb) and truncated L1 insertions used as putative neutral controls in domestic dogs. **b,** Functional enrichment of 1,213 genes located within 50 kb of L1 insertions identified from the pan-genome. Circle area is proportional to the number of input genes annotated to each term. A partial list is

shown; complete enrichment results are provided in Supplementary Tables 7–9. **c,** Tissue-level expression shifts associated with nearby L1 insertions. Each point represents one tissue, with the mean $\log_{10}$(fold-change) across retained L1-adjacent genes plotted against $-\log_{10}$(adjusted P). Genes were filtered by expression level and genotype polymorphism, and tissues with fewer than 50 retained genes were excluded. Colors denote organ systems; the vertical dashed line indicates adjusted P = 0.05. **d,** Gene-level $\log_{10}$(fold-change) distributions across the 20 CNS tissues. Meninges is shown for completeness but was excluded from tissue-level analyses owing to fewer than 50 retained L1-adjacent genes. The vertical red line indicates $\log_{10}$(fold-change) = 0; distributions for non-CNS tissues are shown in Supplementary Fig. 44. **e,** Organ-system-level summary of L1-associated expression shifts. Circle area is proportional to the mean $\log_{10}$(fold-change) across genes in each organ system. P values in **b**, **c** and **e** were adjusted using the Benjamini–Hochberg procedure.

Discussion

The transformation of gray wolves into domestic dogs presents a longstanding challenge in evolutionary biology: how did the key behavioral traits of tameness, including attenuated fearfulness and increased social tolerance, arise over such an exceptionally short timescale without broadly compromising neural and behavioral function? Rather than depending solely on incremental polygenic shifts, the rapid evolution of tameness may also have involved discrete changes in gene regulation. Our findings support a novel model in which TE insertions can act as modular regulatory shortcuts during canine domestication and subsequent human-associated adaptation. By introducing pre-existing regulatory modules into new genomic contexts, TE insertions may create substrates for regulatory exaptation, expanding the routes through which selection can reshape neural gene expression while limiting pleiotropic costs[7,8].

Identifying the genetic substrates underlying these shifts has been difficult because domestic dogs carry a mosaic genomic legacy of bottlenecks, repeated interbreeding, recent breed formation and intense selection. These processes can obscure older adaptive variation and weaken the assumptions of conventional association, allele-frequency or haplotype-based discovery frameworks[2,44,45]. To address this problem, we developed GAMA, an ancestry-aware framework that reconstructs local ancestry across modern canid genomes and tests for lineage-specific enrichment. Applied to 1,929 canids, GAMA recovered a *GPR52*-linked outlier marked by the strongest genome-wide enrichment of the Fertile Crescent–associated ancestry. This finding illustrates how selection-driven distortions in local ancestry can reveal adaptive signals embedded within modern population structure.

At this ancestry-enriched locus, a *Canis*-specific full-length L1-Cf insertion upstream of *GPR52* emerged as the leading candidate cis-regulatory element. Unlike point mutations, which generally modify existing regulatory logic incrementally, a full-length L1 insertion can introduce a discrete 5′ UTR-derived regulatory module in a single insertion event[9,10,20]. The L1-Cf 5′ UTR showed enhancer-like activity in reporter assays, *GPR52* was upregulated in the caudate nucleus in homozygous L1-Cf-bearing

dogs relative to wolf comparison samples, and the neighboring gene *RABGAP1L* showed no comparable difference. Caudate Hi-C detected a local interaction between a *GPR52*-overlapping bin and a mappable region flanking the insertion site. Together, these observations are consistent with a model in which L1-Cf contributed to a reconfigured local regulatory environment at *GPR52*. Although definitive causal attribution will require endogenous perturbation in a native canine context, these convergent ancestry, population-genetic, expression, reporter and chromatin-context data support L1-Cf as a leading regulatory candidate at *GPR52*.

The neurobiology of GPR52 provides a functional bridge between this regulatory change and domestication-relevant behavior. GPR52 modulates cAMP signaling in striatal circuits, including pathways involving DRD2-expressing neurons that participate in aversive learning and avoidance[19,32]. Pharmacological activation of GPR52 in mice reduced looming-induced defensive freezing without broadly altering locomotor or exploratory activity, anxiety-like behavior, social behavior or sensorimotor gating. Crucially, the behavioral profile matches one key dimension of tameness: reduced acute threat reactivity without generalized behavioral impairment. This specificity suggests that domestication-relevant behavioral change may arise in part through precise modulation of neural threat-response thresholds, rather than broad organismal reconfiguration, consistent with evidence that selection for active tameness can proceed without extensive morphological remodeling[46].

The *GPR52*-linked L1-Cf insertion also anchors a broader TE-associated regulatory pattern in the domestic dog lineage. Beyond this locus, full-length L1 insertions showed non-random population-genetic patterns relative to truncated elements, and genes near full-length L1 insertions were enriched for neurodevelopmental, synaptic and neurobehavioral pathways. L1-adjacent genes also showed tissue-specific transcriptional associations, with the strongest signals in central nervous system tissues. These patterns reveal that full-length L1 insertions constitute a pre-existing regulatory reservoir, poised for rapid exaptation under specific selective regimes. This framework also accommodates other L1-adjacent candidates, including metabolic loci such as *SLC5A1*, as additional dimensions of TE-mediated adaptation to human-associated environments[14].

Ancient DNA placed the *GPR52*-linked L1-Cf insertion into temporal and phylogeographic context. L1-Cf was not a dog-specific de novo insertion, but a deeply rooted canid insertion that persisted at low frequency in wolves. In ancient dogs, the earliest detected carrier dated to approximately 6.5 ka, and dropout-adjusted modeling supported a steep rise from low initial frequency. Region-specific analyses inferred the earliest increase in Fertile Crescent–Anatolia–Balkans-related dog populations, followed by later increases across western and northern/eastern Eurasian regions. Although absolute onset times remain uncertain because of uneven archaeological sampling and low-coverage data, the R1-first pattern argues against a recent breed artifact or simple bottleneck-only explanation. Together with the strong association

between L1-Cf and Fertile Crescent–associated local ancestry in modern dogs, these results favor a staged amplification-and-dispersal model in which L1-Cf first rose in Fertile Crescent–Anatolia–Balkans-related dogs before dispersing and rising further in other regions. Thus, the ancient DNA analysis delineates the precise spatiotemporal cascade that transformed this ancestral canid variant into a hallmark of the domestic dog genome.

In conclusion, the L1-Cf locus extends a specific case of canine tameness into a broader principle: rapid phenotypic adaptation can emerge when ancient mobile-element insertions are redeployed as regulatory circuitry. By carrying latent regulatory logic, these elements provide selection with preassembled modules to reshape tissue- and circuit-specific gene expression when ecological pressures shift. In dogs, this paradigm shows how human-associated environments can convert ancient genomic architecture into adaptive behavior through TE-derived regulatory exaptation.

References

1 Larson, G. *et al.* Rethinking dog domestication by integrating genetics, archeology, and biogeography. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 8878-8883, doi:10.1073/pnas.1203005109 (2012).

2 Dutrow, E. V., Serpell, J. A. & Ostrander, E. A. Domestic dog lineages reveal genetic drivers of behavioral diversification. *Cell* **185**, 4737–4755, doi:10.1016/j.cell.2022.11.003 (2022).

3 Darwin, C. *The variation of animals and plants under domestication*. Vol. 2 (Cambridge University Press, 2010).

4 Trut, L., Oskina, I. & Kharlamova, A. Animal evolution during domestication: the domesticated fox as a model. *Bioessays* **31**, 349-360, doi:10.1002/bies.200800070 (2009).

5 Barghi, N., Hermisson, J. & Schlötterer, C. Polygenic adaptation: a unifying framework to understand positive selection. *Nature Reviews Genetics* **21**, 769-781, doi:10.1038/s41576-020-0250-z (2020).

6 Wagner, G. P. & Zhang, J. Z. The pleiotropic structure of the genotype-phenotype map: the evolvability of complex organisms. *Nature Reviews Genetics* **12**, 204-213, doi:10.1038/nrg2949 (2011).

7 Wray, G. A. The evolutionary significance of cis-regulatory mutations. *Nature Reviews Genetics* **8**, 206-216, doi:10.1038/nrg2063 (2007).

8 Carroll, S. B. Evo-devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution. *Cell* **134**, 25-36, doi:10.1016/j.cell.2008.06.030 (2008).

9 Senft, A. D. & Macfarlan, T. S. Transposable elements shape the evolution of mammalian development. *Nature Reviews Genetics* **22**, 691-711, doi:10.1038/s41576-021-00385-1 (2021).

10 Fueyo, R., Judd, J., Feschotte, C. & Wysocka, J. Roles of transposable elements in the regulation of mammalian transcription. *Nat Rev Mol Cell Bio* **23**, 481-497, doi:10.1038/s41580-022-00457-y (2022).

11 Li, X. F. & Liu, N. Advances in understanding LINE-1 regulation and function in the human genome. *Trends in Genetics* **41**, 577-589, doi:10.1016/j.tig.2025.04.011 (2025).

12 Parker, H. G. *et al.* Genomic Analyses Reveal the Influence of Geographic Origin, Migration,

and Hybridization on Modern Dog Breed Development. *Cell Reports* **19**, 697-708, doi:10.1016/j.celrep.2017.03.079 (2017).

13 Meadows, J. R. S. *et al.* Genome sequencing of 2000 canids by the Dog10K consortium advances the understanding of demography, genome function and architecture. *Genome Biol* **24**, 187, doi:10.1186/s13059-023-03023-7 (2023).

14 Axelsson, E. *et al.* The genomic signature of dog domestication reveals adaptation to a starch-rich diet. *Nature* **495**, 360-364, doi:10.1038/nature11837 (2013).

15 Zeder, M. A. Domestication as a model system for niche construction theory. *Evol Ecol* **30**, 325-348, doi:10.1007/s10682-015-9801-8 (2016).

16 Browning, S. R., Waples, R. K. & Browning, B. L. Fast, accurate local ancestry inference with FLARE. *American Journal of Human Genetics* **110**, 326-335, doi:10.1016/j.ajhg.2022.12.010 (2023).

17 Driscoll, C. A., Macdonald, D. W. & O'Brien, S. J. From wild animals to domestic pets, an evolutionary view of domestication. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 9971-9978, doi:10.1073/pnas.0901586106 (2009).

18 Nishiyama, K., Suzuki, H., Maruyama, M., Yoshihara, T. & Ohta, H. Genetic deletion of GPR52 enhances the locomotor-stimulating effect of an adenosine A(2A) receptor antagonist in mice: A potential role of GPR52 in the function of striatopallidal neurons. *Brain Res* **1670**, 24-31, doi:10.1016/j.brainres.2017.05.031 (2017).

19 Komatsu, H. *et al.* Anatomical transcriptome of G protein-coupled receptors leads to the identification of a novel therapeutic candidate GPR52 for psychiatric disorders. *PLoS One* **9**, e90134, doi:10.1371/journal.pone.0090134 (2014).

20 Li, X. *et al.* LINE-1 transcription activates long-range gene expression. *Nat Genet* **56**, 1494-1502, doi:10.1038/s41588-024-01789-5 (2024).

21 Zhang, X. *et al.* The distinct morphological phenotypes of Southeast Asian aborigines are shaped by novel mechanisms for adaptation to tropical rainforests. *Natl Sci Rev* **9**, nwab072, doi:10.1093/nsr/nwab072 (2022).

22 Chi, L. J., Zhang, X. L., Xue, Y. B. & Chen, H. fastHaN: a fast and scalable program for constructing haplotype network for large-sample sequences. *Mol Ecol Resour* **25**, e13829, doi:10.1111/1755-0998.13829 (2025).

23 Chi, L. *et al.* HapNetworkView: a tool for haplotype network exploration and visualization. *BMC Genomics* **26**, 52, doi:10.1186/s12864-025-11206-8 (2025).

24 Sabeti, P. C. *et al.* Detecting recent positive selection in the human genome from haplotype structure. *Nature* **419**, 832-837, doi:10.1038/nature01140 (2002).

25 Zhao, S., Chi, L., Fu, M. & Chen, H. HaploSweep: Detecting and Distinguishing Recent Soft and Hard Selective Sweeps through Haplotype Structure. *Mol Biol Evol* **41**, doi:10.1093/molbev/msae192 (2024).

26 Chen, H., Hey, J. & Slatkin, M. A hidden Markov model for investigating recent positive selection through haplotype structure. *Theor Popul Biol* **99**, 18-30, doi:10.1016/j.tpb.2014.11.001 (2015).

27 Hortenhuber, M. *et al.* The DoGA consortium expression atlas of promoters and genes in 100 canine tissues. *Nat Commun* **15**, 9082, doi:10.1038/s41467-024-52798-1 (2024).

28 Nagel, M., Watanabe, K., Stringer, S., Posthuma, D. & van der Sluis, S. Item-level analyses reveal genetic heterogeneity in neuroticism. *Nat Commun* **9**, 905, doi:10.1038/s41467-018-

03242-8 (2018).

29 Peyrot, W. J. & Price, A. L. Identifying loci with different allele frequencies among cases of eight psychiatric disorders using CC-GWAS. *Nat Genet* **53**, 445-454, doi:10.1038/s41588-021-00787-1 (2021).

30 Pardinas, A. F. *et al.* Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat Genet* **50**, 381-389, doi:10.1038/s41588-018-0059-2 (2018).

31 Ali, S., Wang, P., Murphy, R. E., Allen, J. A. & Zhou, J. Orphan GPR52 as an emerging neurotherapeutic target. *Drug Discov Today* **29**, 103922, doi:10.1016/j.drudis.2024.103922 (2024).

32 Soares-Cunha, C., Coimbra, B., Sousa, N. & Rodrigues, A. J. Reappraising striatal D1- and D2-neurons in reward and aversion. *Neurosci Biobehav Rev* **68**, 370-386, doi:10.1016/j.neubiorev.2016.05.021 (2016).

33 Poulter, S. *et al.* Preclinical pharmacokinetics, metabolism, and disposition of NXE0041178, a novel orally bioavailable agonist of the GPR52 receptor with potential for treatment of schizophrenia and related psychiatric disorders. *Xenobiotica* **55**, 151-166, doi:10.1080/00498254.2025.2501593 (2025).

34 Poulter, S. *et al.* The Identification of GPR52 Agonist HTL0041178, a Potential Therapy for Schizophrenia and Related Psychiatric Disorders. *Acs Med Chem Lett* **14**, 499-505, doi:10.1021/acsmedchemlett.3c00052 (2023).

35 De Franceschi, G., Vivattanasam, T., Saleem, A. B. & Solomon, S. G. Vision Guides Selection of Freeze or Flight Defense Strategies in Mice. *Current Biology* **26**, 2150-2154, doi:10.1016/j.cub.2016.06.006 (2016).

36 Yilmaz, M. & Meister, M. Rapid Innate Defensive Responses of Mice to Looming Visual Stimuli. *Current Biology* **23**, 2011-2015, doi:10.1016/j.cub.2013.08.015 (2013).

37 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).

38 Durand, N. C. *et al.* Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Syst* **3**, 95-98, doi:10.1016/j.cels.2016.07.002 (2016).

39 Rao, S. S. P. *et al.* A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell* **159**, 1665-1680, doi:10.1016/j.cell.2014.11.021 (2014).

40 Gardner, E. J. *et al.* The Mobile Element Locator Tool (MELT): population-scale mobile element discovery and biology. *Genome Research* **27**, 1916-1929, doi:10.1101/gr.218032.116 (2017).

41 Um, J. W. *et al.* Structural basis for LAR-RPTP/Slitrk complex-mediated synaptic adhesion. *Nat Commun* **5**, 5423, doi:10.1038/ncomms6423 (2014).

42 Yim, Y. S. *et al.* Slitrks control excitatory and inhibitory synapse formation with LAR receptor protein tyrosine phosphatases. *Proc Natl Acad Sci U S A* **110**, 4057-4062, doi:10.1073/pnas.1209881110 (2013).

43 Clifford, M. A. *et al.* EphA7 signaling guides cortical dendritic development and spine maturation. *Proc Natl Acad Sci U S A* **111**, 4994-4999, doi:10.1073/pnas.1323793111 (2014).

44 Lin, A. T., Fairbanks, R. A., Barba-Montoya, J., Liu, H. L. & Kistler, L. A legacy of genetic entanglement with wolves shapes modern dogs. *Proc Natl Acad Sci U S A* **122**, e2421768122, doi:10.1073/pnas.2421768122 (2025).

45 Pritchard, J. K., Pickrell, J. K. & Coop, G. The Genetics of Human Adaptation: Hard Sweeps, Soft Sweeps, and Polygenic Adaptation. *Current Biology* **20**, R208-R215, doi:10.1016/j.cub.2009.11.055 (2010).

46 Venkatachalam, B., Biswa, B. B., Nagayama, H. & Koide, T. Association of tameness and sociability but no sign of domestication syndrome in mice selectively bred for active tameness. *Genes Brain and Behavior* **23**, e12887, doi:10.1111/gbb.12887 (2024).

Acknowledgments

This work was supported by National Natural Science Foundation of China (Grant Nos. 32370669, 32400500, and 32388102), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB38040200), the Science and Technology Innovation 2030-Major Project (2022ZD04017), the National Science and Technology Innovation 2030 Major Project of China (2021ZD0203900), the National Key R&D Program of China (2019YFA0707101) and Strategic Priority Research Program of the Chinese Academy of Sciences (XDA0460302). Genomic variation data were produced by the Dog10K Project, an international collaboration to advance canine genetics.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterial.pdf
- SupplementaryTables.xlsx