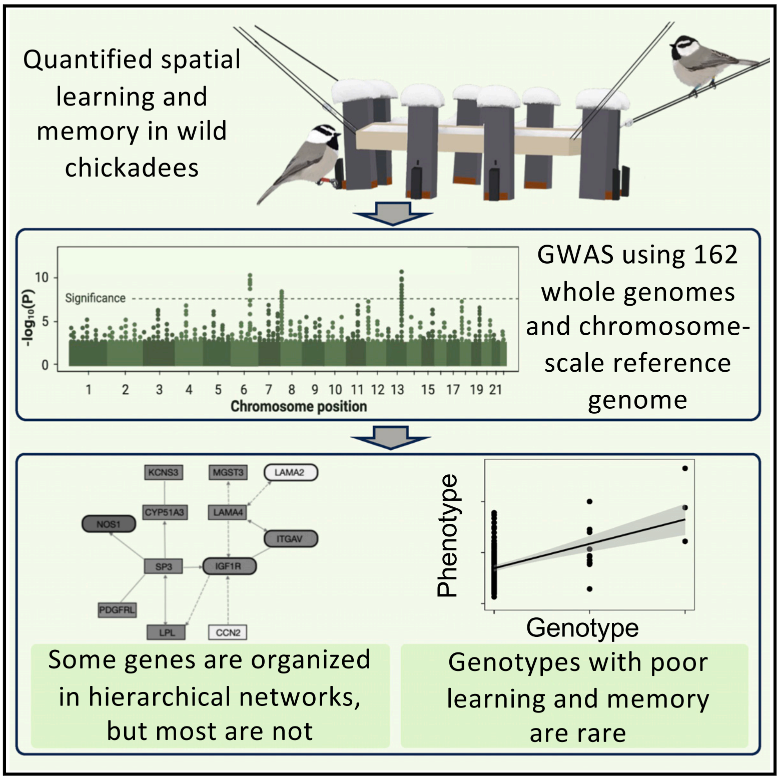
Article

Current Biology

Genes and gene networks underlying spatial cognition in food-caching chickadees

Graphical abstract



Authors

Georgy A. Semenov,
Benjamin R. Sonnenberg,
Carrie L. Branch, ..., Eli S. Bridge,
Vladimir V. Pravosudov, Scott A. Taylor

Correspondence

georgy.semenov@colorado.edu

In brief

Mountain chickadees rely on specialized memory for recovering cached food sources, and variation in this trait has strong effects on fitness. Semenov et al. identify genes and gene networks underlying spatial memory and hypothesize that population variation in memory exists due to a tradeoff between behavioral flexibility and long-term memory.

Highlights

- Nearly one hundred genes have a strong association with spatial memory in chickadees
- Mildly deleterious (possibly regulatory) mutations underly variation in memory
- Multiple developmental mechanisms may contribute to variation in spatial memory
- A large proportion of candidate genes are associated with synaptic plasticity

Semenov et al., 2024, Current Biology *34*, 1930–1939

https://doi.org/10.1016/j.cub.2024.03.058

CellPress

CellPress

Current Biology

Article

Genes and gene networks underlying spatial cognition in food-caching chickadees

Georgy A. Semenov,[1,7,*] Benjamin R. Sonnenberg,[2] Carrie L. Branch,[3] Virginia K. Heinen,[2] Joseph F. Welklin,[2]
Sara R. Padula,[1] Ajay M. Patel,[1] Eli S. Bridge,[4] Vladimir V. Pravosudov,[2,5] and Scott A. Taylor[1,5,6]
[1]Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, Boulder, CO 80309, USA
[2]Department of Biology and Evolution, Ecology Evolution and Conservation Biology Graduate Program, University of Nevada, Reno, NV 89557, USA
[3]Department of Psychology, The University of Western Ontario, London, ON N6A 3K7, Canada
[4]Oklahoma Biological Survey, University of Oklahoma, Norman, OK 73019, USA
[5]Senior author
[6]X (formerly Twitter): @Dr_Scott_Taylor
[7]Lead contact
*Correspondence: georgy.semenov@colorado.edu
https://doi.org/10.1016/j.cub.2024.03.058

SUMMARY

Substantial progress has been made in understanding the genetic architecture of phenotypes involved in a variety of evolutionary processes. Behavioral genetics remains, however, among the least understood. We explore the genetic architecture of spatial cognitive abilities in a wild passerine bird, the mountain chickadee (*Poecile gambeli*). Mountain chickadees cache thousands of seeds in the fall and require specialized spatial memory to recover these caches throughout the winter. We previously showed that variation in spatial cognition has a direct effect on fitness and has a genetic basis. It remains unknown which specific genes and developmental pathways are particularly important for shaping spatial cognition. To further dissect the genetic basis of spatial cognitive abilities, we combine experimental quantification of spatial cognition in wild chickadees with whole-genome sequencing of 162 individuals, a new chromosome-scale reference genome, and species-specific gene annotation. We have identified a set of genes and developmental pathways that play a key role in creating variation in spatial cognition and found that the mechanism shaping cognitive variation is consistent with selection against mildly deleterious non-coding mutations. Although some candidate genes were organized into connected gene networks, about half do not have shared regulation, highlighting that multiple independent developmental or physiological mechanisms contribute to variation in spatial cognitive abilities. A large proportion of the candidate genes we found are associated with synaptic plasticity, an intriguing result that leads to the hypothesis that certain genetic variants create antagonism between behavioral plasticity and long-term memory, each providing distinct benefits depending on ecological context.

INTRODUCTION

A growing collection of studies have made substantial progress in understanding the genetic architecture of phenotypes involved in a variety of evolutionary processes.[1–7] Many of these studies have focused on phenotypic traits with oligogenic genetic architecture (i.e., few genes or large inversions[1,6–9]), but we know relatively little about the genetic underpinnings of traits with more complex genetic architecture. In particular, the study of behavioral genetics in wild organisms is still in its infancy due to the difficulty of experimental quantification of behavioral variation and challenges of disentangling the effects of causal genetic variants from the effects of phenotypic plasticity.[8] Furthermore, studies of genotype-phenotype associations rarely move beyond characterizing the roles of individual genes, resulting in a lack of information about how candidate genes are connected via shared regulation and higher levels of biological organization (e.g., contributions to common developmental pathways).

Due to their inherent complexity and flexibility, behavioral traits often require a high level of integration among diverse developmental and physiological processes. As such, an intriguing question that remains largely unexplored[10] concerns the mechanisms by which selection operates on the genes underlying variation in behavioral traits. One hypothesis is that selection can target genes with upstream positions in biological processes, such as those affecting cascades of biological reactions (e.g., BMP4 in beak formation of birds,[11] ASIP and MC1R in feather patterning,[12] and EDA and morphology in sticklebacks[13]). This hypothesis predicts that, across conspecific individuals, we should observe common patterns of high connectedness among genes due to shared regulation, and such patterns should be identifiable using a genome-wide association study (GWAS). An alternative hypothesis is that selection can

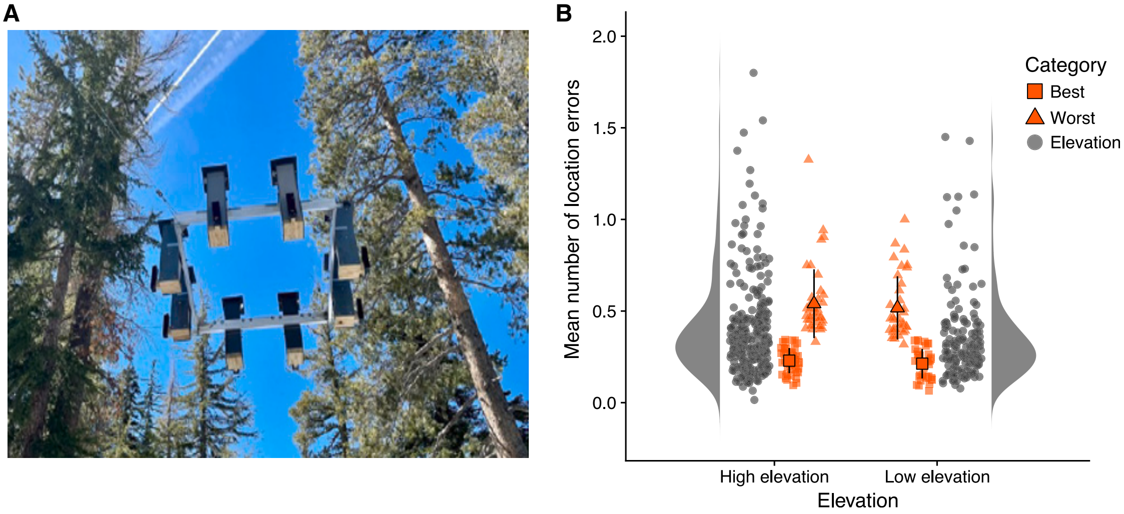


Figure 1. Single smart feeder array and estimates of mean location errors used in this study

(A and B) (A) used for quantifying spatial learning and memory performance and (B) estimates of mean location errors per trial over the 4-day spatial memory testing task. Individual data on all tested birds over the entire study period and their distribution are shown in gray and data points for all birds selected for the analyses here (i.e., best and worst) are shown in orange. Error bars indicate the 95% confidence interval.

operate on a wide array of genes, each with relatively small phenotypic effects on different aspects of development or physiology, leading to a collective phenotypic outcome that is subject to selection, but not characterized by a common network of connected genes (e.g., human height). Although they are fundamental to the field of quantitative genetics, these hypotheses have rarely been evaluated in the context of behavioral genetics in wild animals.[10] Here, we explore these hypotheses by investigating the genetic architecture of spatial cognitive abilities associated with food-caching behavior in wild mountain chickadees (*Poecile gambeli*).

Scatter-caching species, such as mountain chickadees, hide thousands of food items by placing them in unique locations throughout their home range when food is abundant, so that they can recover them when food is scarce. Food caches are essential for overwinter survival and chickadees rely on specialized spatial learning and memory abilities to recover these caches throughout the winter.[14,15] We have previously shown that harsher winter climates impose higher demands on successful cache recovery and spatial cognitive abilities—birds inhabiting higher latitudes (black-capped chickadees, *Poecile attricapillus*) and higher elevations (mountain chickadees, *P. gambeli*) exhibit better spatial learning and memory, have a larger hippocampus, and have more hippocampal neurons compared with their lower latitude and elevation counterparts.[16–20] Focusing on wild mountain chickadees, we have been measuring spatial cognitive abilities in hundreds of individuals in their natural environment annually since 2014 (Figure 1A). We have demonstrated directional natural selection acting on spatial cognitive abilities by showing that individual variation in spatial learning and memory is associated with significant differences in overwinter survival in a cohort of first-year birds.[15] Our most recent genomic work confirmed that spatial cognitive abilities in mountain chickadees have a genetic basis and are heritable.[21]

Although spatial cognition in mountain chickadees appears to be polygenic, we previously identified a smaller set of genes with major phenotypic effects.[21] However, these results were based on a small sample size ($n = 42$), which can produce spurious genetic associations. Nevertheless, this initial study revealed a relative lack of heterozygous loci and alternative homozygote genotypes associated with candidate genes. In some cases, we detected a complete absence of one of the homozygous allelic states for regions of the genome associated with spatial cognition in individuals that performed poorly on the spatial cognitive task. This finding suggested that the mechanism shaping spatial cognition in mountain chickadees may be associated with selection against mildly deleterious mutations. The small sample size in our previous study precluded us from further exploring this possibility or determining the functional state of the mutations (protein-coding vs. non-coding, including regulatory). Finally, we did not attempt to characterize connectedness among the genes underlying variation in spatial cognition with the 42-genome dataset.

In the present study, we use 162 genomes as well as a new chromosome-scale reference genome with species-specific gene annotation to narrow down the list of candidate genes associated with spatial cognition in chickadees (Figure 1B). We explore connectedness between these candidate genes to identify whether they are organized in hierarchical networks, as the aforementioned hypotheses would predict. We assess whether allelic variation related to spatial cognition occurs in protein-coding or non-coding regions of the genome, and we undertake an extensive literature search to identify the molecular mechanisms that may underly population variation in spatial memory. The dataset allows us to confidently identify a set of candidate genes, to explicitly examine the relationship between rare alleles and spatial memory abilities, and to identify whether SNPs associated with differences in spatial cognitive abilities are coding vs. non-coding. We also directly test a hypothesis of functional

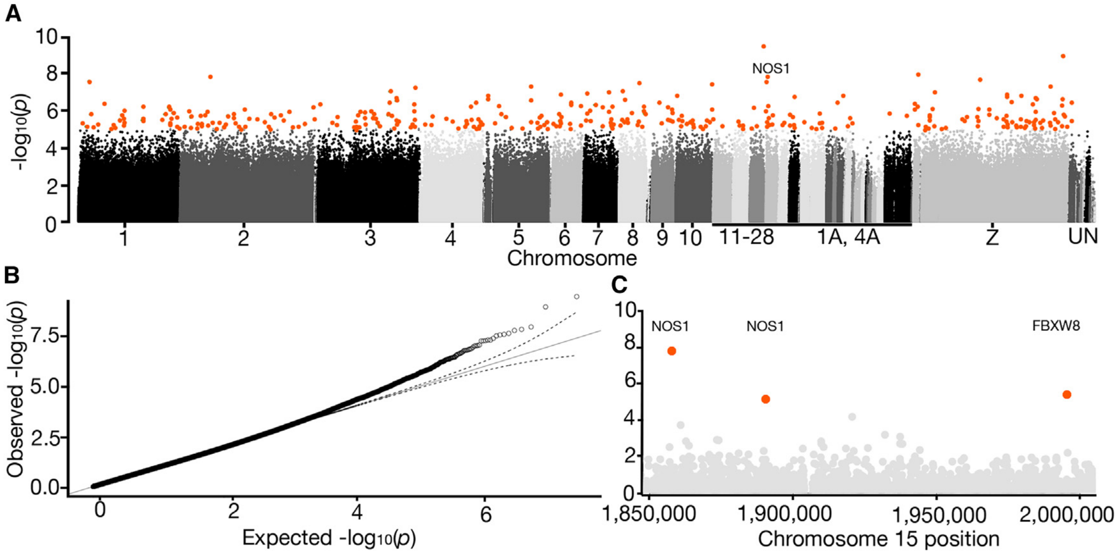


Figure 2. Results of genome-wide association study

(A) Whole-genome associations with outliers above $-\log_{10}(p) = 5$ highlighted in orange ($n = 362$).

(B) Quantile-quantile plot of observed distribution of p values vs. distribution expected under null model of no association signal in GWAS (dashed lines show its 95% confidence interval [CI] for the latter).

(C) An example of clustering of significant associations (orange) in a genomic region containing NOS1 and FBXW8 genes.

hierarchy of the genes associated with spatial cognition (i.e., are they organized in large networks or independent) by using the powerful (and underutilized in evolutionary biology) ingenuity pathway analysis (IPA) tool.

RESULTS

Our chromosome-scale reference assembly produced a total length of ~1.2 Gb organized in 629 scaffolds with an average *N*50 of 73,637,356 bp. The gene annotation recovered 23,949 genes with a total coding region length of 39,134,726 bp. These characteristics represent a substantial improvement with respect to completeness and fragmentation compared with the previously available black-capped chickadee reference genome[22] and comprehensiveness of the previous gene annotation.[21] After applying all filters, our whole-genome dataset included 13,158,377 SNPs for downstream analysis.

The GWAS identified 362 significant genetic associations with the spatial cognitive phenotype in mountain chickadees, greatly exceeding the number expected under false discovery rate (FDR), with the quantile-quantile plot suggesting a breakpoint in overrepresentation of observed associations above $-\log_{10}(p) = 4$ (Figures 2A and 2B; Table S1). Our analysis of average heterozygosity on the Z chromosome separated all individuals into two distinct clusters, 71 females and 91 males. Rerunning GWAS with binary sex as a covariate recovered a similar number of significant associations ($n = 385$, $-\log_{10}(p) = 5$), of which 355 overlapped with the GWAS without sex as a covariate (92%). Our current study ($n = 162$) provided measurable improvement to our previous study ($n = 42$).[21] Consistent with our expectations, the larger dataset produced fewer overall significant associations while increasing the median strength of the association signal (Figure S1).

There were multiple clusters of significant loci in narrow genomic regions across the mountain chickadee genome, supporting a strong biological signal in our genotype-phenotype associations (e.g., Figure 2C). Significantly associated loci ($-\log_{10}(p) \geq 5$, $n = 362$) were found around 97 genes with characterized functions (Table S1), of which 95 occurred in the IPA database (Figure 3). GEMMA Bayesian sparse linear mixed models (BSLMM) analysis revealed that these significant SNPs explain 92% (±12%) of the variance in the spatial cognitive phenotype, whereas a random subset of SNPs accounted for 11% (0%–24%). Further, BSLMM suggested that the posterior number of SNPs in the model most likely explaining trait variance was 69, of which 31 were loci of large effect. The majority of significant SNPs resided in non-coding regions upstream or downstream of a gene reading frame (within 5,000 bp), with the only exception being the CNTNA5 gene (Figure 3), where a SNP resulted in a non-synonymous substitution within the population. Non-coding SNPs within 1,000 bp from the protein-coding part were found near 14 genes (ADAMTS7, ARSI, BUD13, CCN2, CEP55, CNTNAP5, JARID2, KCNS3, LPL, MANSC1, MGST3, PIK3C3, PLEKHF1, and RREB1; Table S1). These SNPs are primary candidate functional regulatory elements.

Results from the IPA found that the genes associated with chickadee spatial cognitive abilities occurred in two large (14–15 genes) highly interconnected networks, one smaller interconnected network (5 genes), and three connected gene pairs (Figure 3). Over half of all cognition-associated genes are not known to be connected to others via a direct or indirect influence of gene expression (Figure 3). The genes recovered from the GWAS are involved in the determination of behavior, cognition, spatial memory, nervous system development and function, as well as neurological disorders (Figure 4; Table S2).

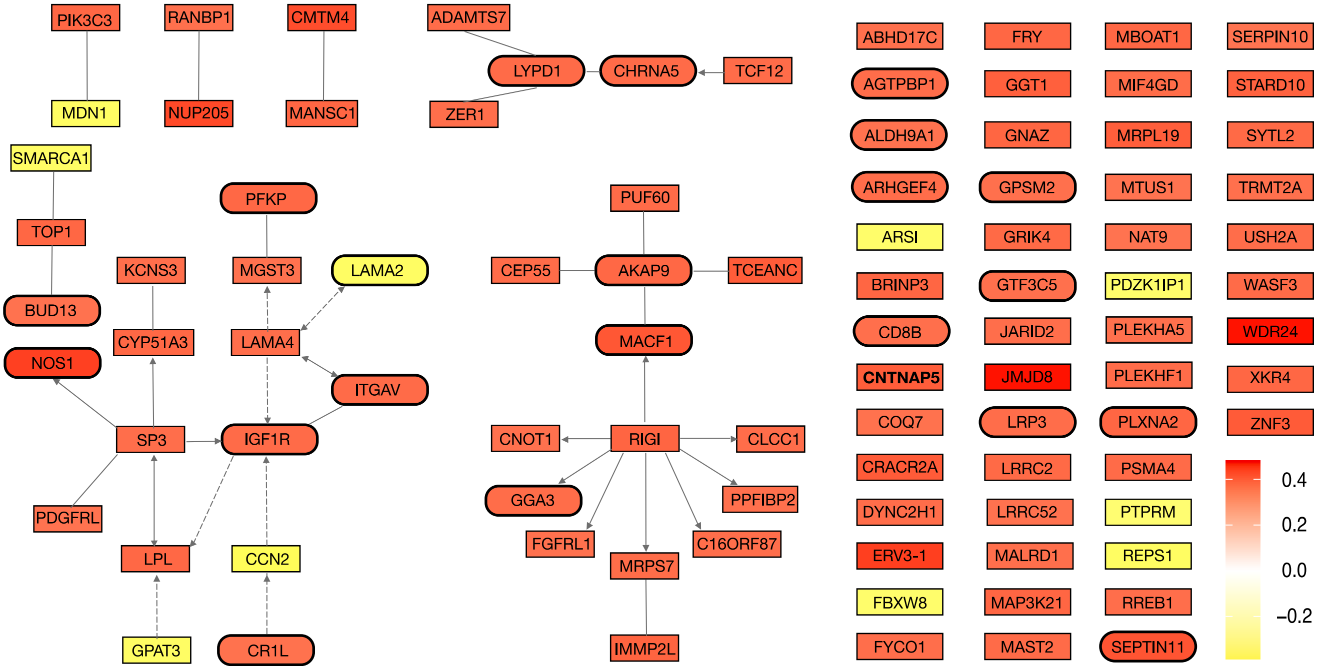


Figure 3. Results of ingenuity pathway analysis showing known direct (solid lines) or indirect (dashed lines) influences on gene expression among the genes recovered in GWAS

Arrows indicate the direction of effect where it is known. Boldface identifies the only gene (CNTNAP5) where an associated SNP was found in the protein-coding region and caused a non-synonymous amino acid substitution. Oval outlines indicate genes associated with synaptic plasticity (see Table S2 for details). Fill color for each gene is scaled to reflect r^2 between genotype and phenotype (see Figure 6 for examples).

Over half of the significantly overrepresented canonical pathways (based on the set of 95 genes) have known connections to neurological functions in the brain (Figure 5). Genes where associated SNPs were particularly strongly correlated with the cognitive phenotype often manifested in a pattern wherein the rare (usually non-reference) allele was homozygous in individuals with "worse" performance (e.g., relatively larger mean number of errors per trial over the task) on the spatial cognitive task (Figure 6).

DISCUSSION

We examined the genetic architecture of spatial cognitive abilities in mountain chickadees using 162 whole genomes

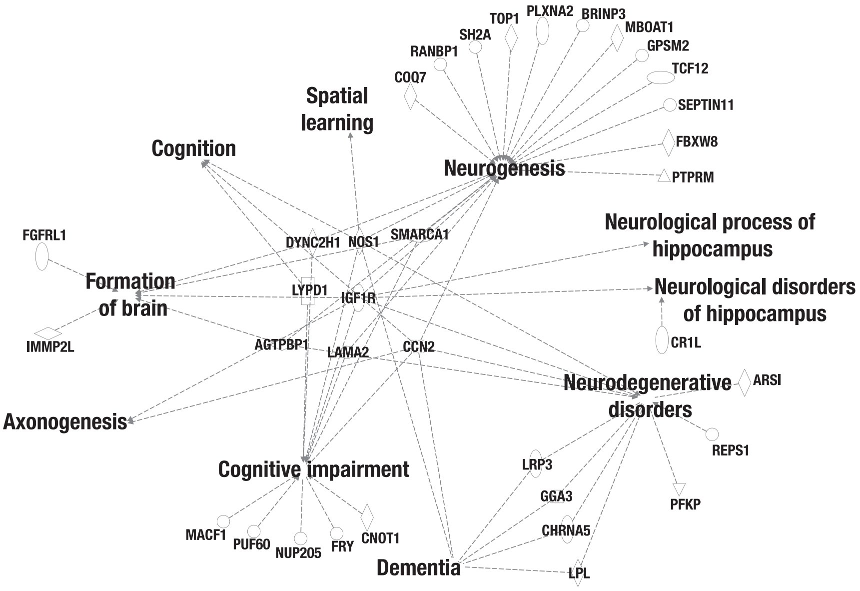


Figure 4. Connections between significant genes identified in GWAS and higher-level functions and phenotypes relevant for nervous system development, functions, and disorders identified using ingenuity pathway analysis

Note that this is only a subset (36 of 97) genes that have known connections to the above categories.

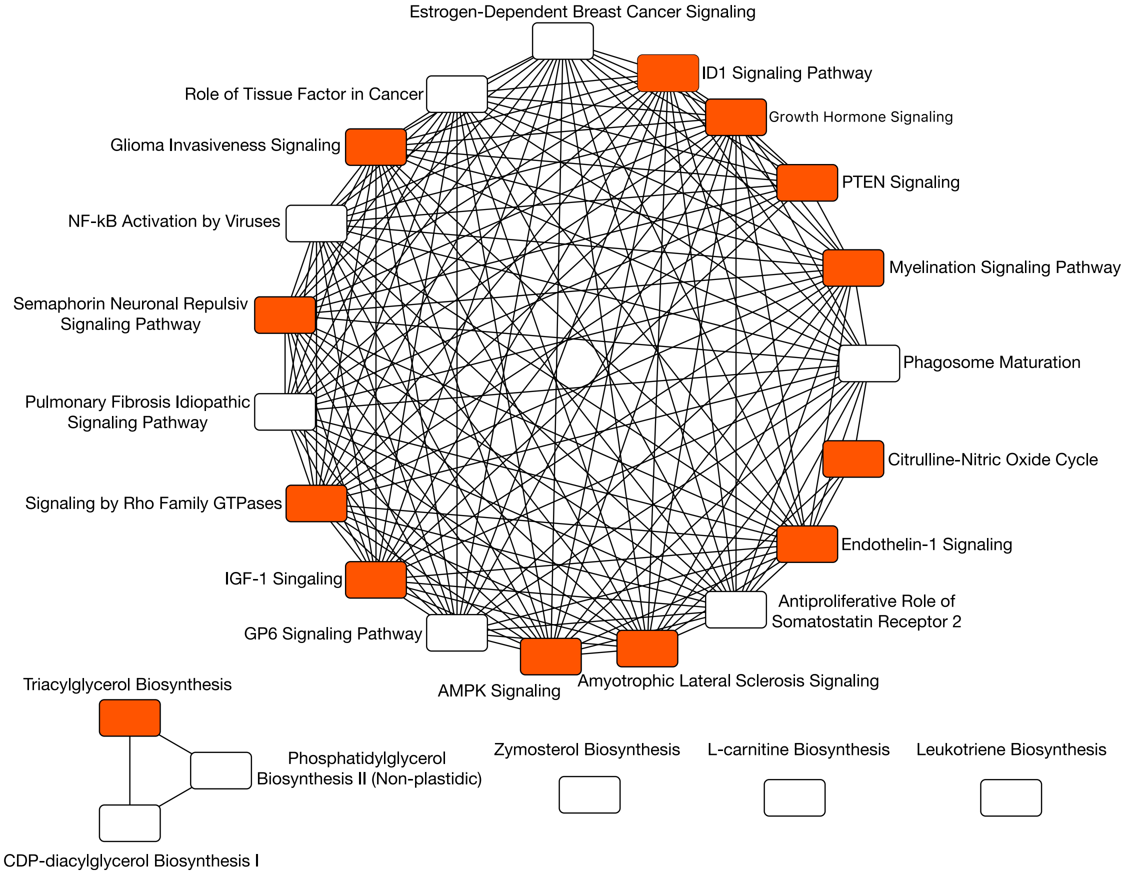


Figure 5. Canonical pathways significantly overrepresented ($p < 0.05$) in the set of 95 genes identified by GWAS and found in the IPA database
Orange-filled rectangles are pathways with connections to neurological functions.

sequenced from a wild, free-living population. Spatial learning and memory are critically important for recovering cached food and natural variation in these cognitive abilities affects mountain chickadee life-time fitness and overwinter survival, particularly in harsh environmental conditions.[15,23] Expanding upon our previous study by using a larger whole-genome dataset, coupled with a new chromosome-scale mountain chickadee reference genome and species-specific gene annotation, we report a strong signal of genotype-phenotype association for spatial learning and memory ability for 97 genes with characterized functions. The associated genes are broadly connected to development and function of the nervous system as well as various neurological disorders. These findings both confirm the validity of our measurements of cognitive abilities and highlight which genes related to cognition may be affected by selection. Although some genes exist within well-characterized gene regulatory networks, about half do not have known or documented connections through shared regulation or previously documented patterns of co-expression. We show that 99% of the significant SNPs reside near, but outside of, protein-coding regions. Fourteen of these are within 1,000 bp of a gene reading frame and likely affect gene expression. Further, we show that, in many cases, rare (non-reference) alleles are associated with poorer spatial cognitive performance in heterozygous individuals or when in the alternative homozygous allelic state (Figure 6). These latter patterns are consistent with what we previously observed using a smaller set of genotyped individuals.[21]

Our current study found ∼8 times fewer associated genes compared with our analysis of 42 individuals analyzed using the same GWAS method.[21] Nonetheless, 15 identical genes were recovered in both studies (CCN2, CNOT1, CYP51A1, EYS, GGA3, JMJD8, MACF1, MBOAT1, NOS1, PDZK1IP1, REPS1, RREB1, SYTL5, TCF12, and USH2A) and 26 additional genes (Table S3) identified in the current study are paralogs or are from the same gene family as those reported in Branch et al.[21] Importantly, our current work used a mountain chickadee reference genome with species-specific annotation, while the previous study used a lower-quality, black-capped chickadee genome and gene annotation evidence from a different species (with slightly different gene nomenclature). As such, the number of overlapping genes should be considered conservative. Including almost 4 times more individuals (compared with Branch et al.[21]) led to a measurable increase in the strength of the association signal, decreasing the number of marginally significant (and potentially spurious) associations, while increasing the median association strength (Figure S1), suggesting improvement in the biological validity of candidate genes underlying variation in spatial learning and memory ability in mountain chickadees.

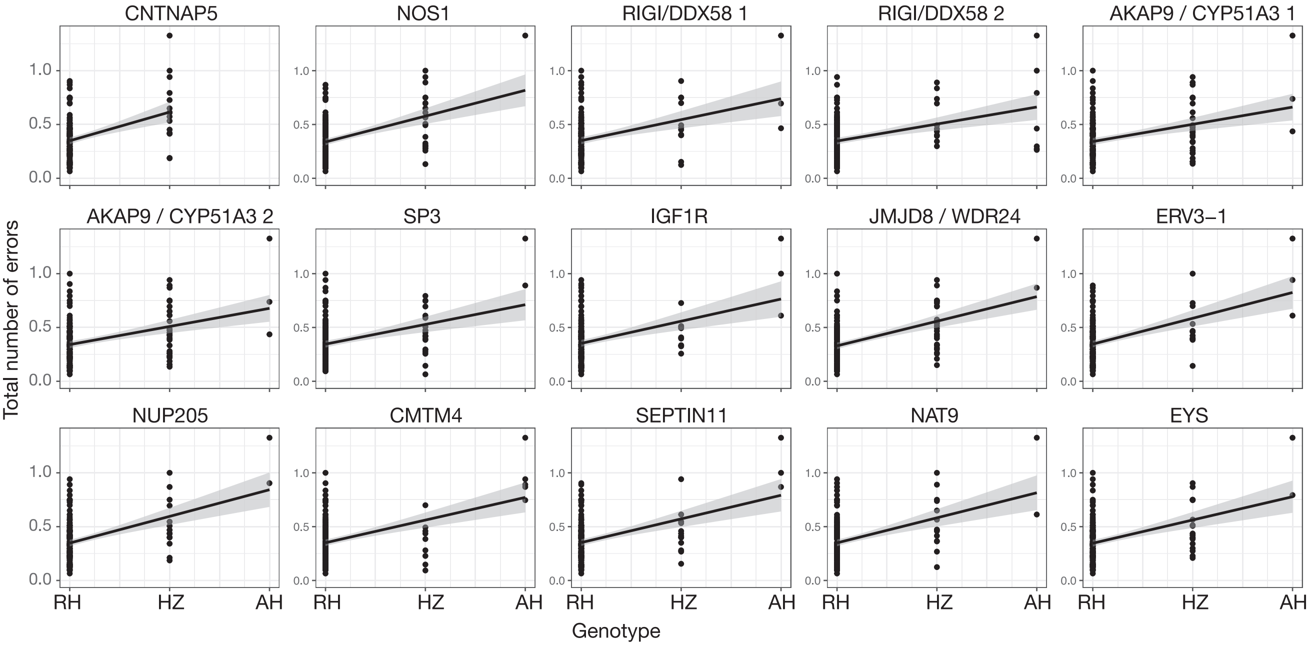


Figure 6. Relationship between genotype and phenotype for select SNPs

CNTNAP5—the only gene with a non-synonymous SNP wherein an alternative homozygote was not found in our population sample. RIGI/DDX58, AKAP/CYP51A3, SP3, and IGF1R—genes with a high upstream regulatory position in developmental processes and central to large networks recovered by IPA. NOS1—gene with high r^2 between genotype and phenotype, and experimentally shown to cause abnormal behavior and relation to behavioral flexibility in mice. Other panels display the genes with the highest r^2 values between genotype and phenotype. RH, HZ, and AH correspond to reference homozygote, heterozygote, and alternative homozygote, respectively. Some of the worst-performing individuals in our dataset had genotypes with alternative homozygotes found in several of the above loci. For example, the individual with worst memory score (#163) had alternative homozygote genotypes for all of these genes, and two more individuals had genotypes including three alternative genotypic states (#L760, genes: JMJD8/WDR24, CMTM4, and SEPTIN11), (#L867, genes: RIGI/DDX58, IGF1R, and SEPTIN11). Lines and gray shadings indicate linear regression and their 95% intervals, respectively.

Compared with the previous study,[21] our current analyses highlight several networks and pathways that may play a key role in generating variation in spatial learning and memory abilities in chickadees. The most significantly overrepresented pathway in our IPA analysis (p = 9.06E−05) was *Signaling by Rho Family GTPases* and included seven genes in our dataset out of 267 known for this pathway (IPA core database). *Rho Family GTPases* have been characterized as critical in neuronal development, survival, and neurodegeneration[24] and are involved in dendritic arborization, spine morphogenesis, growth cone development, and axon guidance. Of particular note, numerous studies connect *Rho Family GTPases* with neurogenesis in the hippocampus (reviewed in Stankiewicz and Linseman[24]) providing an intriguing link between our current study and previous research that showed an association between hippocampal morphology, hippocampal neuron number, adult hippocampal neurogenesis, and spatial cognitive ability in chickadees, including mountain chickadees.[14,19,21] The observation that some of the genes associated with spatial cognitive abilities comprise a part of larger regulatory networks, while others are not interconnected, suggest that there may be many alternative ways (i.e., mutations in different genes with similar end-effects) to produce variation in the spatial cognitive abilities of chickadees. However, several biological pathways (i.e., large gene networks) and genes with upstream regulatory roles in biological processes (Figures 3 and 4) may have particularly prominent functions. We therefore see mixed support regarding how selection operates on spatial cognitive abilities in chickadees—both large-effect genes with upstream positions in development as well as developmentally or physiologically unconnected genes (i.e., possibly small-effect genes or genes with more modular position in development) appear to underlie the trait.

The genetic variants associated with population variation in spatial cognitive function are almost exclusively non-coding and likely play a regulatory role (although we cannot rule out the possibility that some are linked to causal variants). Unlike non-synonymous substitutions, *cis*-regulatory mutations affect gene expression of nearby genes and produce a quantitative rather than qualitative phenotypic effect.[25,26] Such mutations often play a role in fine-tuning phenotypic variation and are more likely to have mild phenotypic effects compared with variants affecting protein structure.[25,26] The majority of genes significantly associated with spatial learning and memory in chickadees appear to be associated with various behavioral diseases and disorders (such as schizophrenia, Alzheimer's disease, and abnormal behaviors, particularly in mice and humans) (Table S2). The observation that heterozygous and, particularly, alternative homozygous individuals for the rare alleles (Figure 6, rare in the studied populations), displayed worse spatial cognitive abilities suggests that some variation in the spatial cognitive abilities of chickadees arises due to mildly deleterious mutations that likely affect gene expression. The only gene with a non-synonymous protein-coding substitution (CNTNAP5) did not occur in an alternative homozygous state in our dataset

(Figure 6), suggesting that the homozygous genotype has a particularly strong negative effect on an individual's survival, at least in the studied ecological context. Variation in CNTNAP5 protein-coding regions is associated with severe behavioral abnormalities in humans (e.g., Pagnamenta et al.[27], and Hongyao et al.[28]), making it a particularly strong candidate as a large-effect gene determining spatial memory variation in mountain chickadees.

The evidence of strong directional selection on spatial cognitive abilities in our high elevation study population of mountain chickadees[15] contrasts with the existence of mutations that produce a worse spatial cognition phenotype. One plausible explanation is that better spatial learning and memory ability is adaptive in some ecological contexts, but not others, and that the degree to which a given rare allele is deleterious is context specific. Such context-specific tradeoffs can be found in other taxa: for example, mutations underlying selective advantages in an antibiotic-rich environment in bacteria can impose fitness costs in an environment that is antibiotic-free.[29,30] Intriguingly, we recently reported that mountain chickadees inhabiting high elevations with harsh environments exhibit lower cognitive flexibility despite showing better performance on the spatial learning and memory task compared with chickadees from lower and milder elevations.[31,32] We tested cognitive flexibility by allowing birds to rapidly learn frequently changing environmental associations.[32,33] The reduced cognitive flexibility that we documented in chickadee that perform well on the spatial learning and memory task is likely a result of proactive memory interference due to stronger memories and a larger memory-load associated with more food caches.[33]

Whether the genes and genetic variants we recovered in the current study are mechanistically connected to this behavioral tradeoff is an outstanding question. Importantly, over 1/5 of the genes (n = 21) identified by GWAS are associated with synaptic plasticity, which is the ability of synapses to alter their strength over time, particularly depending on changes in perceived stimuli. Examples of genes broadly involved in synaptic plasticity include GGA3, LAMA2, PFKP, NOS1, ARHGEF4, CD8B, GTF3C5, and PLXNA2, with some having characterized plasticity mechanisms through regulating GABAergic pathways, neuronal cytoskeleton, cholinergic pathways, astrocytes, amyloid beta plaque, neuronal growth cones, and axonal guidance (e.g., MACF1, AKAP9, IGF1R, ITGAV, CR1L, BUD13, LYPD1, CHRNA5, AGTPB1, ALDH9A1, GPSM2, LRP3, and SEPTIN11; Table S2). Synaptic plasticity may work antagonistically between different forms of memory.[34] For example, mutant transgenic mice (dominant negative for PKA gene regulatory subunit, a gene not found in our study) had improved cognitive flexibility and working memory, while having deficits in long-term reference memory.[34] Of the genes we found in our study, knock out of ARHGEF4, a negative regulator of synaptic plasticity, in lab mice was shown to enhance long-term memory,[35] and LYPD1 are known to influence the filtering of redundant stimuli in mice.[36] Another gene recovered by both our previous and current studies, NOS1 (Figures 3, 4, and 6; Table S2), has a critical role specifically in learning and memory and has been extensively studied for at least three decades.[37] Gene knockout and functional inhibition studies of NOS1 have shown impaired memory and cognitive function in humans, mice, and rats,[37] and NOS1 has also been shown to play an important role in synaptic plasticity in the brains of adult animals, particularly in experience-dependent plasticity.[38] We hypothesize, given our new results, that mutations associated with less effective spatial memory phenotypes in mountain chickadees may be connected to the observed trade-off between memory and behavioral flexibility, particularly through the mechanism of synaptic plasticity.

Despite numerous studies concerning the consolidation of memory, the mechanisms underlying cognitive flexibility, particularly as it relates to reversal learning, remain poorly understood. Adult hippocampal neurogenesis appears to be one such mechanism wherein experimentally induced neurogenesis results in worse memory retention, but better acquisition of new memories, while experimentally reduced neurogenesis (e.g., longer survival of existing neurons) results in better memory retention and worse new memory acquisition.[39,40] We have evidence that better spatial memory in mountain chickadees is negatively associated with performance in a serial reversal task (a metric relevant for cognitive flexibility),[31,32] and we have strong evidence that high elevation birds are less cognitively flexible compared with low elevation birds.[31,32] Our future work will focus on analyzing the genetic architecture of cognitive flexibility to test whether cognitive flexibility and long-term memory in chickadees are encoded by an overlapping set of genes or alternative variants, as would be predicted by the memory-plasticity trade-off hypothesis. Overall, our combined behavioral and genomic data are consistent with the hypothesis that, in harsher winter environments (e.g., higher elevations and latitudes), better spatial learning and memory abilities are particularly important for survival[12] and hence can be expected to result in the reduction of alleles underlying behavioral plasticity. In contrast, across other parts of the mountain chickadee distribution (e.g., at lower elevations and latitudes), increased cognitive flexibility may be more beneficial as milder environments may be more stochastic, requiring birds to learn multiple or changing sources of information.[31,32] This hypothesis will be tested in the future by assessing the genetics of spatial cognition across distributional and elevational ranges and between closely related species in the Paridae family.

Finally, it is interesting that there are no sex-based differences in spatial learning and memory abilities or hippocampal morphology in food-caching chickadees from multiple populations.[17,20,41–43] We found that the vast majority (82%) of the significant GWAS associations are on autosomes and that including sex as a covariate had a minimal effect on GWAS results. This lack of sex-related differences is likely due to strong and equal dependence on spatial cognitive abilities used in food caching for survival in both male and female chickadees. In contrast, in some other species there are large sex-related differences in cognition, which likely evolved due to different evolutionary pressures.[44] For example, in parasitic cowbirds, females that keep track of multiple host nests in which they lay their eggs have better spatial memory and larger hippocampi compared with males that do not nest search.[45–47] In contrast, there are no sex-related differences in spatial cognition in several closely related but not parasitic species. In polygynous meadow voles, males keep track of territories of multiple females and have better spatial learning abilities compared with females, while no sex-related differences have been detected in closely related pine voles.[48]

These contrasting data suggest that different ecological and evolutionary pressures may lead to the evolution of differences in cognition between the sexes. In many species, however, including food-caching chickadees, selection pressures on spatial cognitive abilities appear similar for males and females, leading to the evolution of the same genetic mechanisms underlying cognitive abilities.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
    - Lead contact
    - Materials availability
    - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
    - Study subjects, site, and sampling
- METHOD DETAILS
    - Cognitive testing in the wild
    - DNA extraction, sequencing, variant calling, sexing, reference genome, and annotation
- QUANTIFICATION AND STATISTICAL ANALYSIS
    - Quantification of spatial memory
    - Genome-wide association study (GWAS)
    - Ingenuity pathway analysis
    - Literature search for candidate gene functions
    - Characterization of the non-coding vs. protein-coding state of variants

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2024.03.058.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (NSF) IOS1856181 and IOS2119824 awards to V.V.P. and IOS1928891 and IOS2119825 awards to S.A.T. We thank Carla Cicero (University of California Museum of Vertebrate Zoology) for help with collecting and loaning the tissues used for the reference genome assembly and Matt Carling (University of Wyoming) for help with sampling chickadee tissues for gene annotation (USFWS Migratory Bird Permit MB06336A). We thank Andrew Goodspeed and Etienne Danis (University of Colorado Cancer Center) for providing access to the IPA software.

AUTHOR CONTRIBUTIONS

V.V.P. and S.A.T. conceived and designed the study and obtained funding. B.R.S., C.L.B., V.K.H., J.F.W., and V.V.P. collected all cognitive data and blood samples in the field. V.K.H., J.F.W., and V.V.P. analyzed the cognitive data and selected individuals for the genomic analyses. E.S.B. and V.V.P. co-designed the "smart" RFID feeders, and V.V.P. designed the spatial arrays used to test spatial memory. E.S.B. designed RFID boards and wrote the software. G.A.S. collected tissues in the field for the reference genome annotation, analyzed genomic data, and wrote the manuscript with input from all co-authors. All co-authors approved the final version of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: October 6, 2023
Revised: December 6, 2023
Accepted: March 26, 2024
Published: April 17, 2024

REFERENCES

1. Semenov, G.A., Linck, E., Enbody, E.D., Harris, R.B., Khaydarov, D.R., Alström, P., Anderson, L., and Taylor, S.A. (2021). Asymmetric introgression reveals the genetic architecture of a plumage trait. Nat. Commun. *12*, 1. https://doi.org/10.1038/s41467-020-20314-w.
2. Powell, D.L., Payne, C., Banerjee, S.M., Keegan, M., Bashkirova, E., Cui, R., Andolfatto, P., Rosenthal, G.G., and Schumer, M. (2021). The genetic architecture of variation in the sexually selected sword ornament and its evolution in hybrid populations. Curr. Biol. *31*, 923–935.e11. https://doi.org/10.1016/j.cub.2020.12.049.
3. Morris, J., Navarro, N., Rastas, P., Rawlins, L.D., Sammy, J., Mallet, J., and Dasmahapatra, K.K. (2019). The genetic architecture of adaptation: convergence and pleiotropy in Heliconius wing pattern evolution. Heredity *123*, 138–152. https://doi.org/10.1038/s41437-018-0180-0.
4. Hager, E.R., Harringmeyer, O.S., Wooldridge, T.B., Theingi, S., Gable, J.T., McFadden, S., Neugeboren, B., Turner, K.M., Jensen, J.D., and Hoekstra, H.E. (2022). A chromosomal inversion contributes to divergence in multiple traits between deer mouse ecotypes. Science *377*, 399–405. https://doi.org/10.1126/science.abg0718.
5. Lamichhaney, S., Fan, G., Widemo, F., Gunnarsson, U., Thalmann, D.S., Hoeppner, M.P., Kerje, S., Gustafson, U., Shi, C., Zhang, H., et al. (2016). Structural genomic changes underlie alternative reproductive strategies in the ruff (Philomachus pugnax). Nat. Genet. *48*, 84–88. https://doi.org/10.1038/ng.3430.
6. Funk, E.R., Mason, N.A., Pálsson, S., Albrecht, T., Johnson, J.A., and Taylor, S.A. (2021). A supergene underlies linked variation in color and morphology in a Holarctic songbird. Nat. Commun. *12*, 6833. https://doi.org/10.1038/s41467-021-27173-z.
7. Tuttle, E.M., Bergland, A.O., Korody, M.L., Brewer, M.S., Newhouse, D.J., Minx, P., Stager, M., Betuel, A., Cheviron, Z.A., Warren, W.C., et al. (2016). Divergence and functional degradation of a sex chromosome-like supergene. Curr. Biol. *26*, 344–350. https://doi.org/10.1016/j.cub.2015.11.069.
8. Toews, D.P.L., Taylor, S.A., Vallender, R., Brelsford, A., Butcher, B.G., Messer, P.W., and Lovette, I.J. (2016). Plumage genes and little else distinguish the genomes of hybridizing warblers. Curr. Biol. *26*, 2313–2318. https://doi.org/10.1016/j.cub.2016.06.034.
9. Harringmeyer, O.S., and Hoekstra, H.E. (2022). Chromosomal inversion polymorphisms shape the genomic landscape of deer mice. Nat. Ecol. Evol. *6*, 1965–1979. https://doi.org/10.1038/s41559-022-01890-0.
10. Hoekstra, H.E., and Robinson, G.E. (2022). Behavioral genetics and genomics: Mendel's peas, mice, and bees. Proc. Natl. Acad. Sci. USA *119*, e2122154119. https://doi.org/10.1073/pnas.2122154119.
11. Abzhanov, A., Protas, M., Grant, B.R., Grant, P.R., and Tabin, C.J. (2004). Bmp 4 and morphological variation of beaks in Darwin's finches. Science *305*, 1462–1465. https://doi.org/10.1126/science.1098095.
12. Gluckman, T.L., and Mundy, N.I. (2017). The differential expression of MC1R regulators in dorsal and ventral quail plumages during embryogenesis: Implications for plumage pattern formation. PLoS One *12*, e0174714. https://doi.org/10.1371/journal.pone.0174714.
13. Jones, F.C., Grabherr, M.G., Chan, Y.F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M.C., White, S., et al. (2012). The genomic basis of adaptive evolution in threespine sticklebacks, Nature *7392*, 55–61.
14. Pravosudov, V.V., Roth, I.I., and T.C.. (2013). Cognitive ecology of food hoarding: the evolution of spatial memory and the hippocampus. Annu.

Rev. Ecol. Evol. Syst. *44*, 173–193. https://doi.org/10.1146/annurev-ecolsys-110512-135904.

15. Sonnenberg, B.R., Branch, C.L., Pitera, A.M., Bridge, E., and Pravosudov, V.V. (2019). Natural selection and spatial cognition in wild food-caching mountain chickadees. Curr. Biol. *29*, 670–676.e3. https://doi.org/10.1016/j.cub.2019.01.006.
16. Roth, I.I., Pravosudov, T.C., and V.V.. (2009). Hippocampal volume and neuron numbers increase along a gradient of environmental harshness – a large-scale comparison. Proc. Biol Soc. Lond. B *276*, 401–405. https://doi.org/10.1098/rspb.2008.1184.
17. Roth, I.I., T.C., LaDage, L.D., and Pravosudov, V.V. (2011). Variation in hippocampal morphology along an environmental gradient: controlling for the effect of day length. Proc. Biol. Sci. *278*, 2662–2667. https://doi.org/10.1098/rspb.2010.2585.
18. Roth, I.I., T.C., LaDage, L.D., Freas, C., and Pravosudov, V.V. (2012). Variation in memory and the hippocampus across populations from different climates: a common garden approach. Proc. Biol. Sci. *279*, 402–410. https://doi.org/10.1098/rspb.2011.1020.
19. Freas, C.A., LaDage, L.D., Roth, I.I., T.C., and Pravosudov, V.V. (2012). Elevation related differences in memory and the hippocampus in food-caching mountain chickadees. Anim. Behav. *84*, 121–127. https://doi.org/10.1016/j.anbehav.2012.04.018.
20. Freas, C.A., Bingman, K., LaDage, L.D., and Pravosudov, V.V. (2013). Untangling elevation-related differences in the hippocampus in food-caching mountain chickadees: the effect of a uniform captive environment. Brain Behav. Evol. *82*, 199–209. https://doi.org/10.1159/000355503.
21. Branch, C.L., Semenov, G.A., Wagner, D.N., Sonnenberg, B.R., Pitera, A.M., Bridge, E.S., Taylor, S.A., and Pravosudov, V.V. (2022). The genetic basis of spatial cognitive variation in a food-caching bird. Curr. Biol. *32*, 210–219.e4. https://doi.org/10.1016/j.cub.2021.10.036.
22. Wagner, D.N., Curry, R.L., Chen, N., Lovette, I.J., and Taylor, S.A. (2020). Genomic regions underlying metabolic and neuronal signaling pathways are temporally consistent outliers in a moving avian hybrid zone. Evolution *74*, 1498–1513. https://doi.org/10.1111/evo.13970.
23. Branch, C.L., Pitera, A.M., Kozlovsky, D.Y., Sonnenberg, B.R., Benedict, L.M., and Pravosudov, V.V. (2019). Elevation-related differences in the age structure of breeding birds suggest stronger selection at harsher elevations. Behav. Ecol. Sociobiol. *73*, 143. https://doi.org/10.1007/s00265-019-2750-4.
24. Stankiewicz, T.R., and Linseman, D.A. (2014). Rho family GTPases: key players in neuronal development, neuronal survival, and neurodegeneration. Front. Cell. Neurosci. *8*, 314. https://doi.org/10.3389/fncel.2014.00314.
25. Wittkopp, P.J., and Kalay, G. (2011). Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. Nat. Rev. Genet. *13*, 59–69. https://doi.org/10.1038/nrg3095.
26. Wray, G.A. (2007). The evolutionary significance of cis-regulatory mutations. Nat. Rev. Genet. *8*, 206–216. https://doi.org/10.1038/nrg2063.
27. Pagnamenta, A.T., Bacchelli, E., de Jonge, M.V., Mirza, G., Scerri, T.S., Minopoli, F., Chiocchetti, A., Ludwig, K.U., Hoffmann, P., Paracchini, S., et al. (2010). Characterization of a family with rare deletions in CNTNAP5 and DOCK4 suggests novel risk loci for autism and dyslexia. Biol. Psychiatry *68*, 320–328. https://doi.org/10.1016/j.biopsych.2010.02.002.
28. Hongyao, H.E., Chun, J.I., Xiaoyan, G., Fangfang, L., Jing, Z., Lin, Z., Pengxiang, Z., and Zengchun, L. (2023). Associative gene networks reveal novel candidates important for ADHD and dyslexia comorbidity. BMC Med. Genomics *16*, 208. https://doi.org/10.1186/s12920-023-01502-1.
29. Tanaka, M.M., and Valckenborgh, F. (2011). Escaping an evolutionary lobster trap: drug resistance and compensatory mutation in a fluctuating environment. Evolution *65*, 1376–1387. https://doi.org/10.1111/j.1558-5646.2011.01223.x.
30. Lenski, R.E. (1998). Bacterial evolution and the cost of antibiotic resistance. Int. Microbiol *1*, 265–270. https://doi.org/10.2436/IM.V1I4.27.
31. Croston, R., Branch, C.L., Pitera, A., Kozlovsky, D.Y., Parchman, T.L., Bridge, E.S., and Pravosudov, V.V. (2017). Predictably harsh environment is associated with reduced cognitive flexibility in wild mountain chickadees. Anim. Behav. *123*, 139–149. https://doi.org/10.1016/j.anbehav.2016.10.004.
32. Benedict, L.M., Heinen, V.K., Sonnenberg, B.R., Bridge, E.S., and Pravosudov, V.V. (2023). Learning predictably alternating food locations across days in a food-caching bird. Anim. Behav. *196*, 55–81. https://doi.org/10.1016/j.anbehav.2022.11.005.
33. Tello-Ramos, M.C., Branch, C.L., Kozlovsky, D.Y., Pitera, A.M., and Pravosudov, V.V. (2019). Spatial memory and cognitive flexibility trade-offs: to be or not to be flexible, that is the question. Anim. Behav. *147*, 129–136. https://doi.org/10.1016/j.anbehav.2018.02.019.
34. Malleret, G., Alarcon, J.M., Martel, G., Takizawa, S., Vronskaya, S., Yin, D., Chen, I.Z., Kandel, E.R., and Shumyatsky, G.P. (2010). Bidirectional regulation of hippocampal long-term synaptic plasticity and its influence on opposing forms of memory. J. Neurosci. *30*, 3813–3825. https://doi.org/10.1523/JNEUROSCI.1330-09.2010.
35. Yoo, K.S., Lee, K., Lee, Y.S., Oh, W.J., and Kim, H.K. (2020). Rho Guanine Nucleotide Exchange Factor 4 (Arhgef4) deficiency enhances spatial and object recognition memory. Exp. Neurobiol. *29*, 334–343. https://doi.org/10.5607/en20049.
36. Sherafat, Y., Chen, E., Lallai, V., Bautista, M., Fowler, J.P., Chen, Y.C., Miwa, J., and Fowler, C.D. (2021). Differential expression patterns of Lynx proteins and involvement of Lynx1 in repulse inhibition. Front. Behav. Neurosci. *15*, 703748. https://doi.org/10.3389/fnbeh.2021.703748.
37. Weitzdoerfer, R., Hoeger, H., Engidawork, E., Engelmann, M., Singewald, N., Lubec, G., and Lubec, B. (2004). Neuronal nitric oxide synthase knock-out mice show impaired cognitive performance. Nitric Oxide *10*, 130–140. https://doi.org/10.1016/j.niox.2004.03.007.
38. Donohoe, G., Walters, J., Morris, D.W., Quinn, E.M., Judge, R., Norton, N., Giegling, I., Hartmann, A.M., Möller, H.J., Muglia, P., et al. (2009). Influence of NOS1 on verbal intelligence and working memory in both patients with schizophrenia and healthy control subjects. Arch. Gen. Psychiatry *66*, 1045–1054. https://doi.org/10.1001/archgenpsychiatry.2009.139.
39. Akers, K.G., Martinez-Canabal, A., Restivo, L., Yiu, A.P., de Cristofaro, A., Hsiang, H.L., Wheeler, A.L., Guskjolen, A., Niibori, Y., Shoji, H., et al. (2014). Hippocampal neurogenesis regulates forgetting during adulthood and infancy. Science *344*, 598–602. https://doi.org/10.1126/science.1248903.
40. Frankland, P.W., Köhler, S., and Josselyn, S.A. (2013). Hippocampal neurogenesis and forgetting. Trends Neurosci. *36*, 497–503. https://doi.org/10.1016/j.tins.2013.05.002.
41. Petersen, K., and Sherry, D.F. (1996). No sex difference occurs in hippocampus, food-storing, or memory for food caches in black-capped chickadees. Behav. Brain Res. *79*, 15–22. https://doi.org/10.1016/0166-4328(95)00257-x.
42. Pravosudov, V.V., and Clayton, N.S. (2002). A test of the adaptive specialization hypothesis: population differences in caching, memory and the hippocampus in black-capped chickadees (Poecile atricapilla). Behav. Neurosci. *116*, 515–522. https://doi.org/10.1037/0735-7044.116.4.515.
43. Branch, C.L., Sonnenberg, B.R., Pitera, A.M., Benedict, L.M., Kozlovsky, D.Y., Bridge, E., and Pravosudov, V.V. (2020). Testing the greater male variability phenomenon: male mountain chickadees exhibit larger variation in reversal learning compared to females. Proc. Biol. Soc. L *287*, 20200895. https://doi.org/10.1098/rspb.2020.0895.
44. Clint, E.K., Sober, E., Garland, T., Jr., and Rhodes, J.S. (2012). Male superiority in spatial navigation: adaptation or side effect? Q. Rev. Biol. *87*, 289–313. https://doi.org/10.1086/668168.
45. Guigueno, M.F., Snow, D.A., MacDougall-Shackleton, S.A., and Sherry, D.F. (2014). Female cowbirds have more accurate spatial memory than males. Biol. Lett. *10*, 20140026. https://doi.org/10.1098/rsbl.2014.0026.
46. Guigueno, M.F., MacDougall-Shackleton, S.A., and Sherry, D.F. (2015). Sex differences in spatial memory in brown-headed cowbirds: males

outperform females on a touchscreen task. PLoS ONE *10*, e0128302. https://doi.org/10.1371/journal.pone.0128302.

47. Sherry, D.F., and Guigueno, M.F. (2019). Cognition and the brain of brood parasitic cowbirds. Integr. Zool. *14*, 145–157. https://doi.org/10.1111/1749-4877.12312.
48. Jacobs, L.F., Gaulin, S.J., Sherry, D.F., and Hoffman, G.E. (1990). Evolution of spatial cognition: sex-specific patterns of spatial behavior predict hippocampal size. Proc. Natl. Acad. Sci. USA *87*, 6349–6352. https://doi.org/10.1073/pnas.87.16.6349.
49. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
50. Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc.
51. Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM preprint at arXiv.
52. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map Format and SAMtools. Bioinformatics *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.
53. Broad Institute (2018). Picard Toolkit. GitHub. http://broadinstitute.github.io/picard/.
54. Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al. (2011). The Variant Call Format and VCFtools. Bioinformatics *27*, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330.
55. Putnam, N.H., O'Connell, B.L., Stites, J.C., Rice, B.J., Blanchette, M., Calef, R., Troll, C.J., Fields, A., Hartley, P.D., Sugnet, C.W., et al. (2016). Chromosome-scale shotgun assembly using an in vitro method for long-range linkage. Genome Res. *26*, 342–350. https://doi.org/10.1101/gr.193474.115.
56. Marçais, G., Delcher, A.L., Phillippy, A.M., Coston, R., Salzberg, S.L., and Zimin, A. (2018). MUMmer4: A fast and versatile genome alignment system. PLoS Comput. Biol. *14*, e1005944. https://doi.org/10.1371/journal.pcbi.1005944.
57. Flynn, J.M., Hubley, R., Goubert, C., and Smit, A.F. (2020). RepeatModeler2 for automated genomic discovery of transposable element families. Proc. Natl. Acad. Sci. USA *117*, 9451–9457. https://doi.org/10.1073/pnas.1921046117.
58. Levitsky, V.G. (2004). RECON: a program for prediction of nucleosome formation potential. Nucleic Acids Res. *32*, W346–W349. https://doi.org/10.1093/nar/gkh482.
59. Price, A.L., Jones, N.C., and Pevzner, P.A. (2005). De novo identification of repeat families in large genomes. Bioinformatics *21* (*Supplement 1*), i351–i358. https://doi.org/10.1093/bioinformatics/bti1018.
60. Smit, A.F., Hubley, R., and Green, P. (2013–2015). RepeatMasker Open-4.0. http://www.repeatmasker.org.
61. Stanke, M., and Morgenstern, B. (2005). AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. Nucleic Acids Res. *33*, W465–W467. https://doi.org/10.1093/nar/gki458.
62. Korf, I. (2004). Gene finding in novel Genomes. BMC Bioinformatics *5*, 59. https://doi.org/10.1186/1471-2105-5-59.
63. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.
64. Chan, P.P., and Lowe, T.M. (2019). tRNAscan-SE: Searching for tRNA Genes in Genomic Sequences. Methods Mol. Biol. 1–14. https://doi.org/10.1007/978-1-4939-9173-0_1.
65. Zhou, X., and Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. Nat. Genet. *44*, 821–824. https://doi.org/10.1038/ng.2310.
66. Krämer, A., Green, J., Pollard, J., Jr., and Tugendreich, S. (2014). Causal analysis approaches in Ingenuity Pathway Analysis. Bioinformatics *30*, 523–530. https://doi.org/10.1093/bioinformatics/btt703.
67. Nelson, C.W., Moncla, L.H., and Hughes, A.L. (2015). SNPGenie: estimating evolutionary parameters to detect natural selection using pooled next-generation sequencing data. Bioinformatics *31*, 3709–3711. https://doi.org/10.1093/bioinformatics/btv449.
68. Croston, R., Kozlovsky, D.Y., Branch, C.L., Parchman, T.L., Bridge, E.S., and Pravosudov, V.V. (2016). Individual variation in spatial memory performance in wild mountain chickadees from different elevations. Anim. Behav. *111*, 225–234. https://doi.org/10.1016/j.anbehav.2015.10.015.
69. Tello-Ramos, M.C., Branch, C.L., Pitera, A.M., Kozlovsky, D.Y., Bridge, E.S., and Pravosudov, V.V. (2018). Memory in wild mountain chickadees from different elevations: comparing first year birds with older survivors. Anim. Behav. *137*, 149–160. https://doi.org/10.1016/j.anbehav.2017.12.019.
70. Bridge, E.S., Wilhelm, J., Pandit, M.M., Moreno, A., Curry, C.M., Pearson, T.D., Proppe, D.S., Holwerda, C., Eadie, J.M., Stair, T.F., et al. (2019). An Arduino-based RFID Platform for Animal Research. Front. Ecol. Evol. *7*, 257. https://doi.org/10.3389/fevo.2019.00257.
71. Benedict, L.M., Heinen, V.K., Sonnenberg, B.R., Pitera, A.M., Bridge, E.S., and Pravosudov, V.V. (2023). Food-caching chickadees do not exhibit directional bias when learning a spatial task. Behav. Ecol. Sociobiol. *77*, 4. https://doi.org/10.1007/s00265-022-03275-6.
72. Heinen, V.K., Pitera, A.M., Sonnenberg, B.R., Branch, C.L., Benedict, L.M., Welklin, J.F., Whitenack, L.E., Bridge, E.S., and Pravosudov, V.V. (2013). Food-caching chickadees with specialized spatial cognition do not use scrounging as a stable strategy when learning a spatial task. Proc. R. Soc. Lond. B *290*, 20230900.
73. Heinen, V., Pitera, A.M., Sonnenberg, B.R., Branch, C.L., Bridge, E., and Pravosudov, V.V. (2021). Specialized spatial cognition is associated with reduced cognitive senescence in a food-caching bird. Proc. R. Soc. Lond. B *288*, 20203180.
74. Poplin, R., Ruano-Rubio, V., DePristo, M.A., Fennell, T.J., Carneiro, M.O., Van der Auwera, G.A., Kling, D.E., Gauthier, L.D., Levy-Moonshine, A., Roazen, D., et al. (2017). Detailed description of HaplotypeCaller; best reference for germline joint calling preprint at bioRxiv.
75. Benedict, L.M., Pitera, A.M., Branch, C.L., Sonnenberg, B.R., Heinen, V.K., Bridge, E.S., and Pravosudov, V.V. (2021). Information maintenance of food sources is associated with environment, spatial cognition and age in a food-caching bird. Anim. Behav. *182*, 153–172. https://doi.org/10.1016/j.anbehav.2021.10.009.
76. https://advaitabio.com/science/pathway-analysis-vs-gene-set-analysis/.
77. Pennacchio, L.A., Bickmore, W., Dean, A., Nobrega, M.A., and Bejerano, G. (2013). Enhancers: five essential questions. Nat. Rev. Genet. *14*, 288–295. https://doi.org/10.1038/nrg3458.
78. Jo, B.S., and Choi, S.S. (2015). Introns: The Functional Benefits of Introns in Genomes. Genomics Inform. *13*, 112–118. https://doi.org/10.5808/GI.2015.13.4.112.
79. Le, N.Q.K., Yapp, E.K.Y., Nagasundaram, N., and Yeh, H.Y. (2019). Classifying promoters by interpreting the hidden information of DNA sequences via deep learning and combination of continuous FastText N-Grams. Front. Bioeng. Biotechnol. *7*, 305. https://doi.org/10.3389/fbioe.2019.00305.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Blood samples preserved in Queens lysis buffer	Pravosudov lab	https://chickadeecognition.com
Deposited data		
Raw read data, reference genome assembly and annotation	This paper	https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1084804/
Software and algorithms		
Trimmomatic 0.39	Li et al.[49]	http://www.usadellab.org/cms/?page=trimmomatic
FastQC 0.11.7	Andrews[50]	http://www.bioinformatics.babraham.ac.uk/projects/fastqc
Bwa mem v.0.7.17-r1188	Li[51]	https://github.com/lh3/bwa
Samtools v. 1.3.1	Li et al.[52]	http://www.htslib.org
Picard-tools v.2.8.1	Picard Toolkit[53]	http://broadinstitute.github.io/picard/
VCFtools v.0.1.15	Danecek et al.[54]	https://vcftools.github.io/index.html
HiRise	Putnam et al.[55]	https://bio.tools/hirise
MUMmer v.4.0	Marçais et al.[56]	https://mummer.sourceforge.net
RepeatModeler v.2.0.1	Flynn et al.[57]	https://www.repeatmasker.org/RepeatModeler/
RECON v.1.08	Levitsky[58]	http://eddylab.org/software/recon/
RepeatScout v.1.0.6	Price et al.[59]	https://biocontainer-doc.readthedocs.io/en/latest/source/repeatscout/repeatscout.html
RepeatMasker v.4.1.0	Smit et al.[60]	http://www.repeatmasker.org
AUGUSTUS v.2.5.5	Stanke et al.[61]	https://github.com/Gaius-Augustus/Augustus/blob/master/docs/ABOUT.md
SNAP v. 2006-07-28	Korf[62]	https://github.com/KorfLab/SNAP
STAR v.2.7	Dobin et al.[63]	https://github.com/alexdobin/STAR
tRNAscan-SE v. 2.05	Chan et al.[64]	http://lowelab.ucsc.edu/tRNAscan-SE/
GEMMA v.0.98	Zhou et al.[65]	https://github.com/genetics-statistics/GEMMA
Ingenuity Pathway Analysis	Krämer et al.[66]	https://digitalinsights.qiagen.com/IPA
SNPgenie v.2019.10.31	Nelson et al.[67]	https://github.com/chasewnelson/SNPGenie

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Georgy Semenov (georgy.semenov@colorado.edu).

Materials availability

This study did not generate new unique reagents

Data and code availability

- Raw cognitive testing data for this study is available as supplemental data file included with this manuscript (Data S1). Raw paired whole genome sequencing reads, reference genome and annotation are available through NCIB accession number PRJNA1084804.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Study subjects, site, and sampling

We study food-caching mountain chickadees in the Sierra Nevada mountains in Sagehen Experimental Forest (Sagehen Creek Field Station, University of California Berkeley; 39.432°, -120.237°). We have been banding chickadees with uniquely coded Passive Integrated Transponder leg bands (*i.e.,* PIT tags) since 2014 at two montane elevations. The high elevation (ca. 2300-2500 m) sites are associated with severe winter conditions, including frequent snowstorms and deeper and longer lasting snow cover; whereas low elevation (ca. 1965 - 2070 m) sites exhibit milder winter conditions.[68,69] We trap and band chickadees during fall and winter using mist nets around multiple permanent feeders and in the nest during the breeding season. Upon capture, we collect a blood sample from each bird (ca. 100 uL) from the brachial vein. Collected blood is stored in Queens lysis buffer at 2°C until extraction.

The study was approved by the University of Nevada Reno Institutional Animal Care and Use Committee (Protocol 00818, 00046 and 00603) and was in accordance with California Department of Fish and Wildlife Permit D-0011776516-4. For banding efforts, we followed U.S. Federal Bird Banding Permit 22878.

METHOD DETAILS

Cognitive testing in the wild

We measure spatial cognitive abilities in chickadees using "smart" spatial arrays of Radio Frequency Identification (RFID)-enabled food dispensers or feeders. Each array consists of 8 feeders positioned equidistantly on a square frame (1.2 x 1.2 m) and raised 3+ m above the ground using metal cables and pulleys attached to four trees.[31,68] We have deployed 4 spatial arrays, 2 at each elevation site, ca. 1.5 km apart with almost no overlap in birds using each array. Each feeder has a perch with an embedded RFID antenna that is connected to a customized Arduino circuit board with a built in RFID reader.[70] Hence, individual birds that visit a feeder can be identified by means of their PIT-tag leg bands. Moreover, each feeder can be programmed with a list of individual bird IDs that determines which individuals can access food via a motorized door. Feeders also record the ID and time of an individual landing on the perch regardless of whether food is made available.[15,32,70] During annual testing (January-March), we first maintain each array with all doors open, so birds find food and habituate to the feeders in the array. Next, we program all feeders to keep the doors closed, but to open for any bird with a PIT-tag, allowing birds to habituate to the door opening mechanism. Finally, for cognitive testing, we program the feeders so each bird can only get food from a single feeder within each array. Birds are assigned pseudo-randomly across all 8 feeders, avoiding assigning the feeders used most frequently during habituation periods.[15,32,68,70] For both habituation and testing, feeders are filled with black-oil sunflower seeds, and chickadees only collect one sunflower seed from the feeder per visit (they can only hold one seed in their beak) and fly to the nearby trees to either eat or cache it before returning to the array.

We measure cognitive performance by calculating the number of location errors (*i.e.*, number of unrewarding feeders) a bird makes during each trial. A trial begins when a bird visits any feeder in the array and ends with a visit to the correct, rewarding feeder (the bird's assigned feeder). The spatial learning and memory task lasts four days during each year of testing. The vast majority of birds at our field site (>100 individuals each year and ca. 90-95% of all birds detected at the study sites) participate in cognitive testing. Our test design does not require any manipulation by a bird to open the feeder – all it has to do is to land on the perch and the door either opens so the bird can get food or stays closed.[31,68] If the door stays closed, birds start moving around the array until they land on the feeder that provides food.[32] Our task is a standard test of spatial associative learning and memory and we have previously shown that chickadees use spatial cognitive abilities to learn this task.[68] It has been well documented that performance in a spatial associative learning task reflects learning and memory used for cache recovery.[14] Considering that only one feeder is rewarding, and all other feeders are non-rewarding, birds are expected to learn the rewarding location over multiple trials. Even though chickadees are social, they learn this task individually and do not follow their flock mates to learn and remember the location of the rewarding feeder in the array and their performance on the task is not affected by social learning e.g.[71,72] Furthermore, male and female chickadees do not differ in spatial learning and memory ability or in hippocampal morphology.[17,20,41–43] Spatial learning and memory abilities measured in our task do not change significantly across years or between the first and the second year of life.[15] Finally, cognitive performance does not show senescence across the birds' natural life span.[73]

DNA extraction, sequencing, variant calling, sexing, reference genome, and annotation

We extracted DNA using the QIAGEN DNeasy Blood and Tissue protocol and quantified it using an Invitrogen Qubit 3.0 fluorometer (Invitrogen, Carlsbad CA). Whole genome library preparation was performed using a Nextera XT DNA Library Preparation Kit following standard protocol. Libraries were pooled and sequenced on an Illumina NovaSEQ 6000 (paired end, 150 base pairs) at approximately 15x coverage at the University of Colorado Cancer Center Genomics and Microarray Core Facility. We trimmed low quality reads and removed Illumina adapters using Trimmomatic 0.39[49] and performed quality control on trimmed sequence files using FastQC 0.11.7.[50] Next, bwa mem v.0.7.17-r1188[51] was used to align reads to the chromosome-scale mountain chickadee reference genome (see below). Bam files were sorted, duplicates were marked, and files were indexed using samtools v. 1.3.1[52] and picard-tools v.2.8.1.[53] Variants were called using HaplotypeCaller from GATK v.4.2.[74] The resulting gvcfs were merged using CombineGVCFs and genotyped with GenotypeGVCFs followed by VariantFiltration using GATK-recommended filters. We used VCFtools v.0.1.15[54] to remove indels, keep only biallelic SNPs with minor allele frequency above 5%, minQ>20, min-meanDP>4, max-meanDP<75 and max-missing=0.75.

A new chromosome-scale reference genome and gene annotation of the mountain chickadee were generated by Dovetail Genomics. We used flash-frozen tissues from heart, liver, muscle, kidney, brain, and eyes of a mountain chickadee from California (specimen MVZ:Bird:193498 [https://arctos.database.museum/guid/MVZ:Bird:193498] collected 10/4/2021 at 41.318°, -120.929° and stored at the University of California Museum of Vertebrate Zoology) from which DNA was extracted using a QIAGEN DNeasy Blood and Tissue kit. For the reference genome, PacBio CCS reads and Dovetail OmniC reads were used as input to HiC-Hifiasm with default parameters. This process produced one *de novo* assembly for each haplotype. For Omni-C libraries, chromatin in the nucleus was fixed with formaldehyde. Fixed chromatin was digested with DNase I and then extracted, chromatin ends were repaired and ligated to a biotinylated bridge adapter followed by proximity ligation of adapter-containing ends. After proximity ligation, crosslinks were reversed, and the DNA was purified. Purified DNA was treated to remove biotin that was not internal to ligated fragments. Sequencing libraries were generated using NEBNext Ultra enzymes and Illumina-compatible adapters. Biotin-containing fragments were isolated using streptavidin beads before PCR enrichment of each library. The library was sequenced on an Illumina HiSeqX platform to produce $\sim$ 30x sequence coverage. The input *de novo* assembly and Dovetail OmniC library reads were used as input data for HiRise, a software pipeline designed specifically for using proximity ligation data to scaffold genome assemblies.[55] Dovetail OmniC library sequences were aligned to the draft input assembly using bwa v.0.7.17-r1188.[51] The separations of Dovetail OmniC read pairs mapped within draft scaffolds were analyzed by HiRise to produce a likelihood model for genomic distance between read pairs, and the model was used to identify and break putative mis-joints, to score prospective joins, and make joins above a threshold. We next used MUMmer v.4.0[56] to establish the synteny between the mountain chickadee reference genome and a reference assembly of the zebra finch (*Taeniopygia guttata*) (https://www.ncbi.nlm.nih.gov/assembly/GCA_009859065.1/) for plotting purposes. We ran MUMmer for each zebra finch chromosome separately using default setting except increasing –maxgap to 1000 to account for structural rearrangements.

To produce a species-specific annotation, we sampled breast muscle, liver, brain, lung, gonads, and eye tissue from a male mountain chickadee collected in Wyoming (9/05/2022, 41.170°, -106.070°, study skin is pending deposition to the University of Wyoming Museum of Vertebrates). Tissues were preserved in the field in RNAlater (QIAGEN). Total RNA was extracted using the RNeasy Plus Kit (QIAGEN) following standard protocol and quantified using Qubit RNA Assay and TapeStation 4200. Prior to library prep, we performed DNase treatment followed by AMPure bead clean up and FastSelect HMR rRNA (QIAGEN) depletion. Library preparation was done with the NEBNext Ultra II RNA Library Prep Kit following manufacturer protocols. The libraries were sequenced on the NovaSeq6000 platform in 2 x 150 bp configuration. Repeat families found in the genome assemblies of mountain chickadee were identified *de novo* and classified using the software package RepeatModeler v.2.0.1.[57] RepeatModeler depends on the programs RECON v.1.08[58] and RepeatScout v.1.0.6[59] for the *de novo* identification of repeats within the genome. The custom repeat library obtained from RepeatModeler were used to discover, identify, and mask the repeats in the assembly file using RepeatMasker v.4.1.0.[60] Coding sequences from collared flycatcher (*Ficedula albicollis*), great tit (*Parus major*), and zebra finch (*Taeniopygia guttata*) were used to train the initial *ab initio* model for mountain chickadee using the AUGUSTUS software v.2.5.5.[61] Six rounds of prediction optimization were done with the software package provided by AUGUSTUS. The same coding sequences were also used to train a separate *ab initio* model for mountain chickadee using SNAP v. 2006-07-28.[62] RNAseq reads were mapped onto the genome using the STAR aligner software v.2.7[63] and intron hints generated with the bam2hints tools within the AUGUSTUS software. MAKER, SNAP, and AUGUSTUS (with intron-exon boundary hints provided from RNA-Seq) were then used to predict for genes in the repeat-masked reference genome. To help guide the prediction process, Swiss-Prot peptide sequences from the UniProt database were downloaded and used in conjunction with the protein sequences from collared flycatcher, great tit, and zebra finch to generate peptide evidence in the Maker pipeline. Only genes that were predicted by both SNAP and AUGUSTUS software were retained in the final gene sets. To help assess the quality of the gene prediction, AED scores were generated for each of the predicted genes as part of the MAKER pipeline. Genes were further characterized for their putative function by performing a BLAST search of the peptide sequences against the UniProt database. tRNA were predicted using the software tRNAscan-SE v. 2.05.[64]

Sexing of chickadees using morphology in the field is often not possible outside of breeding season. We therefore estimated mean heterozygosity for loci on the Z chromosome to determine the sex for each individual. Because female birds are heterogametic (i.e., have a single copy of the Z chromosome), males and females form two distinct clusters, with reduced heterozygosity in females.

QUANTIFICATION AND STATISTICAL ANALYSIS

Quantification of spatial memory

To evaluate individual differences in spatial cognitive performance, we used the mean number of location errors per trial over the entire four-day spatial cognitive task, following our previous work documenting natural selection and genetic differences associated with spatial cognitive performance using this metric.[13,19,68] This metric captures improvements in learning and memory performance across multiple trials. As an individual learns across successive trials, it makes fewer and fewer errors each trial. The mean number of errors per trial across multiple trials reflects such learning – a lower mean number of errors per trial across multiple trials reflects better learning and memory whereas a larger mean number of errors per trial reflects worse learning and memory.[15] Importantly, all birds learn the task00BEthe mean number of location errors per trial was significantly better than it would be expected by chance (i.e., random sampling of the feeders)[15,68] - but there was large individual variation and we have previously shown that this metric accurately reflects learning and memory performance across multiple trials.[15,31,68] Most critically, we have shown that this metric of spatial learning and memory is ecologically relevant because individual variation in this metric is associated with significant differences in survival,[15] parental investment,[23] foraging decisions[75] and performance in a serial reversal learning task.[32]

Genome-wide association study (GWAS)

We performed GWAS on a complete matrix of 13,158,377 SNP loci and a continuous measure of individual spatial learning and memory abilities00BEthe mean number of location errors per trial over the entire spatial learning task in 162 individuals using GEMMA v.0.98.[65] To estimate the effects of individual SNPs we ran Linear Mixed-effects Models (LMM, -lmm 1 option of GEMMA) while supplying a relatedness matrix (the -gk 1option of GEMMA) as a covariate. We focused on significance threshold levels above 1.00E-5 -$\log_{10}(P) = 5$. We associated SNPs with genes based on our whole genome annotation, extending gene coordinates by 5000 base pairs above and below coding sequence to account for non-coding (possibly regulatory) variation (see also Table S1). To evaluate the effects of sex on our GWAS results we also ran LMM while supplying sex as a binary covariate. To assess if including ~4 times more individuals in a GWAS compared to our previous study[21] produced a measurable improvement in the strength of association signal, we ran LMM with the above settings using the subset of individuals from our previous study[21](see also Figure S1; Table S3). To estimate heritability of spatial memory, we used Bayesian Sparse Linear Mixed Models (BSLMM) of GEMMA. We ran four BSLMM chains with a burn-in of five million steps and a subsequent 20 million MCMC steps sampling every 1000 iterations on a subset of SNPs with high association signal in LMM (P<=1.00E-05, n=362). To test for robustness of BSLMM explanatory power, we used a subset of SNPs with low LMM signal (P>1.00E-01) randomly sampled across the genome to produce the sample size as in the GWAS-associated subset (n=362) and re-ran the analysis with the same settings.

Ingenuity pathway analysis

To characterize the organization of individual GWAS-associated genes in developmental processes and regulatory pathways we used the Ingenuity Pathway Analysis (IPA) software by QIAGEN (Release 2022-11-27, QIAGEN Inc., https://digitalinsights.qiagen.com/IPA[66]). Unlike methods that use unstructured gene lists for assessing overrepresentation via gene ontology categories, pathway analysis tools (such as IPA) can establish casual hierarchical relationships from gene lists and provide insights into biological mechanism, not just pattern.[76] The IPA database in particular, combines insight from over 100,000 published datasets and over seven million findings to evaluate patterns of gene co-expression, hierarchical gene expression relationships, and connections between individual genes and a phenotype of interest, providing a unique, yet underutilized, tool for studies in evolutionary biology. To assess overrepresentation of GWAS-associated genes in higher-level pathways, we used Ingenuity Knowledge Base analysis with default settings (Direct gene effects and Indirect effects through one other gene). To establish relationships between individual GWAS-associated genes and their organization in gene regulatory networks, we used IPA Path Explorer while accounting for direct and indirect gene relationships and default analysis settings. We further used the IPA Core database to assess connection of GWAS-associated genes to phenotypes and diseases of interest.

Literature search for candidate gene functions

To better characterize genes in the context of this study, we performed a literature search for the 97 candidate genes identified as significant outliers from GWAS. We used Google Scholar and limited our queries to the "gene name" and any of the following: "synapse", "brain", "neuron", "cognition", "plasticity", "synaptic plasticity" and "spatial cognition". We additionally performed these queries with aliases, to ensure a more thorough search. The search was limited to the first 3 pages, to standardize search rate, and abstracts were read to determine how robustly the focal gene's function was investigated and how relevant the paper was to spatial cognition. Papers that fit these criteria were read and selected for addition to Table S2 for two main reasons. The first, and most mechanistic, is if a knockout experiment corroborated the association between a gene and the spatial cognition phenotype. The second is if a paper experimentally investigated the function of the gene or related genes, particularly in the context of neurophysiology or spatial cognition. The papers were thoroughly read and summarized to mechanistically corroborate associations between a given candidate gene and the spatial cognition phenotype. The genes were further classified into categories (synaptic plasticity, related to synaptic plasticity, brain development, brain maintenance, and uncategorized) based on major functions according to the literature search. For instance, if a knock-out study of the focal gene resulted in a microcephaly phenotype, the gene was categorized as "brain development" because the brain did not develop correctly when the focal gene was manipulated. Genes were labelled as uncategorized (in the context of spatial cognition) if gene function was unrelated or incidental to the brain (i.e., tumorigenesis).

Characterization of the non-coding vs. protein-coding state of variants

We used SNPgenie v.2019.10.31[67] to characterize the functional state of nucleotide substitutions. We first manually output sequence and annotation files for each GWAS-associated genic region. We used additional *perl* scripts provided with the program to convert antisense strand genes to reverse complements (see SNPgenie documentation for details). We then ran the snpgenie.pl script to assess the population state of SNPs (synonymous vs nonsynonymous for protein-coding sequences, or non-coding). While characterizing the functional state of non-coding variants (i.e., their correspondence to promoters, enhancers, intronic regions, *cis* or -*trans*-regulatory elements) is challenging (e.g.,[77]) and is beyond the scope of this study, we classified significant GWAS associations into those in close proximity to coding regions (i.e., within 1000 bp) and those further away (i.e., 1001-5000bp, Table S1). The set of variants that are potential promoters (i.e., within 1000bp of the reading frame) or are found in introns are particularly strong candidates that might affect alternative splicing or gene expression.[78,79]