OXFORD JOURNAL
OF STUDENT
SCHOLARSHIP

Are the Genes Associated with Human Cognitive Function Conserved Across Species that Exhibit Fluid Reasoning and Spatial Cognition, and What Does the Pattern of Conservation Reveal About the Evolution of Intelligence?

Vedant Binani
vedantbinani2205@gmail.com

ABSTRACT

The study seeks to determine whether genes associated with human cognitive function have been conserved in other species, and what the pattern of conservation reveals about the evolution of intelligence. The study uses 350 genes obtained from the intersection of the Davies (2018) and Savage (2018) Genome-Wide Association Studies related to general cognitive function. It then tracks their orthologs across nine other species, spanning more than 600 million years of evolutionary divergence, and finds that human cognitive genes were significantly more conserved than the genome average in every species tested. The paper has tested whether the human cognitive genes overlap with independently compiled gene sets of the mouse, fruit fly, and mountain chickadee. This was conducted using hypergeometric tests against a background of 23,262 human protein-coding genes. The research has found that the human, mouse, and fly cognitive genes overlapped much more than the genome average. However, the chickadee gene set shared no genes with the other three sets and the absence of overlap was not statistically significant. The gene ontology enrichment analysis identified 56 pathways that were independently enriched across all three species, a convergence that permutation testing confirmed exceeds chance. However, functional annotation showed that only 6.2% of conserved mouse orthologs and 3.5% of conserved fly orthologs have documented cognitive roles, showcasing that sequence conservation does not necessarily imply functional conservation. The results of the study support the theory that an ancient molecular toolkit for learning and memory is shared across most animals. However, the selection for cognitive ability across different lineages can act on different genes.

***Keywords*:** Cognitive genomics; Ortholog conservation; Comparative genomics; GWAS; Conserved toolkit mode

1) INTRODUCTION

Humans have transformed nearly every ecosystem on Earth, not through physical superiority but rather their superior intellect. This includes abilities such as planning ahead, solving unfamiliar problems, communicating complex ideas and transmitting knowledge across generations. These abilities raise the evolutionary question of whether the genetic architecture of human cognition is unique to the species, specifically assembled over 600 million years of evolution, or built on an older foundation shared with the rest of the animal kingdom.

Answering this matters beyond evolutionary biology, because many cognition-associated genes overlap with those implicated in cognitive and psychiatric disorders, so identifying which are conserved and functionally shared across model organisms can direct mechanistic and therapeutic research.

It is important to be specific about which cognitive abilities are being studied, as intelligence has multiple different definitions in the current literature. This paper uses the Cattell-Horn-Carroll (CHC) model of cognitive abilities (Cattell, 1963; Carroll, 1993). This is the dominant theoretical framework in modern behavioural genetics and psychometrics. The CHC model is a three-level hierarchy. At the top is general intelligence (g), which captures the positive correlations between performance on almost all cognitive tests (Spearman,1904). Below g are broad abilities. These include fluid reasoning (Gf), crystallised intelligence (Gc), visual-spatial processing (Gv), working memory (Gwm), and processing speed (Gs).

This study focuses on fluid reasoning and spatial cognition. Fluid reasoning is the ability to solve novel problems without relying on prior knowledge. It includes inductive, deductive, and quantitative reasoning, e.g., solving a new puzzle. Spatial cognition is the ability to mentally manipulate shapes, sizes, locations, and spatial relationships. It includes mental rotation, spatial scanning, and navigation. e.g., navigating a complex terrain. Both have substantial heritability, and both are relevant to comparative work because different species exhibit observable variation in tasks associated with these abilities.

A Genome-Wide Association Study (GWAS) scans entire genomes to identify genetic variants, such as single-nucleotide polymorphisms (SNPs), that occur more frequently in individuals who express a specific trait. Modern GWAS has identified hundreds of genes that are associated with cognitive function. Comparative work has suggested that the molecular basis of learning and memory is conserved across the animal kingdom. However, few other studies have systematically mapped human cognitive GWAS genes onto the tree of life at the genetic level, and most have been unidirectional, using human genes. This study takes 350 human cognitive genes and tracks their orthologs across nine species. It then tests overlap with independently compiled cognitive gene sets from three other species. The mountain chickadee data enables a bidirectional analysis

One clarification is essential at the outset, because it defines what the results can and cannot claim. The human gene set used is derived from GWAS of general cognitive function (g) rather than from GWAS of

fluid reasoning or spatial cognition specifically, because no adequately powered GWAS of either trait yet exists. The study therefore uses (g) as a genetic proxy for these two domains. This is defensible because fluid reasoning is the single strongest contributor to g and the two are highly genetically correlated ($r_g \approx 0.88–1.0$) , so most of the genetic basis of fluid reasoning is captured within (g). For spatial cognition the proxy is weaker and only partial: although spatial ability is highly heritable (King, 2019) and shares much of its genetic variance with g (genetic correlation ≈ 0.6), a portion of that variance is specific to spatial ability and independent of g (Rimfeld, 2017). A g-based gene set therefore captures the shared component of spatial cognition but will miss some spatial-specific genes. This is a central reason the study is designed bidirectionally: because g is an incomplete proxy for spatial cognition and no human spatial-cognition GWAS exists, the mountain chickadee spatial-cognition GWAS is used to approach spatial ability directly from the animal side rather than only through g. The gene set should thus be read as the shared genetic core of general cognition rather than as markers unique to fluid reasoning or spatial ability, a distinction carried through the interpretation and revisited in the limitations.

This paper supports the conserved toolkit framework, which holds that the molecular machinery of learning and memory is widely shared across animals and that cognitive evolution proceeds primarily through regulatory rather than protein-coding changes. However, it argues that the current comparative literature overinterprets the pathway-level convergence and often treats ortholog conservation as functional conservation. This study tests both concerns rather than assuming them. The literature review establishes the theoretical frameworks relevant to the paper and reviews the existing evidence. The methodology section describes the data sources, statistical methods, and analytical layers used. The findings and discussion section presents the results and then interprets them. The conclusion then summarises the paper's findings and discusses potential directions for future research.

2) LITERATURE REVIEW

2.1) Conserved Toolkit Model

The ‘conserved toolkit’ model is a divergent deployment model of cognitive evolution. It posits that the core molecular machinery of learning and memory- N-methyl-D-aspartate (NMDA) receptors, the cAMP-PKA-CREB signalling cascade, the postsynaptic density scaffold, and associated synaptic plasticity genes, have been deeply conserved since before the divergence of vertebrates and invertebrates. There are 3 major sources of evidence for this. The first one is Eric Kandel’s Nobel Prize-winning work on the sea slug Aplysia. This established the cAMP-PKA-CREB pathway as the molecular substrate of long-term memory (Kandel, 2001). The second is work on fruit flies that found that classical memory mutants dunce, rutabaga, and amnesiac converge on the same cAMP pathway (Davis, 2005). The third is Nithianantharajah (2013). This study used identical touchscreen paradigms in mice and humans to demonstrate that DLG2 (a postsynaptic density scaffold gene) is required for complex cognitive function in both species.

The deployment side of the model states that species-specific cognitive abilities arise primarily from gene regulation and the modification of existing genes, rather than the creation of new coding sequences. One source of evidence for this is the Human Accelerated Regions (HARs). These are conserved genomic loci that have undergone rapid nucleotide substitutions following divergence from chimpanzees. They can act as enhancers which can control when and where specific genes are turned on or off during development. According to Cui (2023), nearly half of all HARs function as neural enhancers in the brain. One HAR-dense gene is FOXP2. This contains two unique amino acid changes in humans compared to chimpanzees. It encodes a transcription factor that regulates hundreds of other genes. It is associated with complex speech in humans, producing different transcriptional target networks (Konopka, 2009). Human-specific gene duplications also contribute: SRGAP2C slows neuronal maturation, raising synaptic density in the cortex (Charrier, 2012), and ARHGAP11B drives neocortical expansion by increasing neural stem-cell production (Florio, 2015).The second source of evidence comes from the New Caledonian Crow, where Dussex (2021) found subtle genetic changes associated with tool use, with the selected genes related to eye structure and bill morphology rather than cognition, again supporting regulatory fine-tuning over coding-sequence change.

This study will test the conserved toolkit portion through conservation and overlap analysis, and the divergent deployment portion through the comparison of human and chickadee cognitive gene sets.

2.2) Human Cognitive Genetics

An individual's intelligence is influenced by a combination of genetic and environmental factors. Epigenetics connects genes to the external environment through tags that attach to genes, making them more or less active. From the genetic standpoint, intelligence appears to be highly heritable. Classical twin studies have consistently estimated the heritability of general intelligence to be 50% in childhood, rising to 70%-80% in adulthood (Bouchard and McGue, 1981; Plomin and Deary, 2015). The increase in heritability with age is known as the Wilson Effect Paradox; genetic influences on cognition become more important as people age. The heritability of spatial cognition has been estimated to be approximately 61% in a meta-analysis of 42 twin studies (King, 2019). These heritability estimates are the starting point for all subsequent molecular work. They led to the identification of cognitive genes responsible for intelligence.

The first attempts to identify specific genes for intelligence were candidate gene studies. Scientists selected genes thought to contribute to intelligence, based on their roles in neurotransmission or brain development and tested whether variants affected cognitive test performance in small samples. These studies failed after the arrival of GWAS. Chabris (2012) tested 12 of the most widely cited candidate-gene associations for intelligence in over 10,000 individuals and failed to replicate any at genome-wide significance. Similar replication failures throughout the behavioural genetics literature led to an agreement that most reported candidate gene associations were false positives driven by low statistical power, publication bias, and flexible analytic choices in small samples. The lessons from this failure are

relevant as large sample sizes are essential for reliable gene discovery, and individual genes associated with cognition can have very small effects. Therefore, the intersection-based gene selection strategy used in this study retains only genes that appear in both Davies and Savage's (2018) GWAS studies to reduce the probability of false positives.

2.3) The Importance of GWAS

Modern human cognitive genetics is based on multiple large-scale GWAS, with larger sample sizes resulting in greater statistical power and accuracy. There are several examples of important GWAS studies relevant to this paper. Sniekers (2017, N = 78,308) identified 52 genes across 18 loci. Savage (2018, N = 269,867) expanded this to 1,016 genes and 205 loci, with the MEF2C, SHANK3, and SATB2 genes among the key hits. Davies (2018, N = 300,486) identified 709 genes and 148 loci from the UK Biobank, testing verbal-numerical reasoning. Hill (2019) used MTAG to combine GWAS of intelligence and educational attainment. They found 538 genes, with neurogenesis as the most enriched pathway ($p = 5.59 \times 10^{-10}$). Lee (2018, N ≈ 1.1 million) found 1,271 genome-wide significant SNPs. The enriched pathways across all these studies are very consistent: neurogenesis, synaptic plasticity, dendrite development, myelination, and ion channel activity.

2.4) Connecting Neurobiology and Cognitive Genetics

These genetic frameworks map onto specific brain regions. Jung and Haier (2007) identified the lateral prefrontal and parietal cortices as the neural basis of fluid reasoning in their Parieto-Frontal Integration Theory. Spatial Cognition is predominantly regulated by the CA1 region of the hippocampus and the parahippocampal and retrosplenial cortices. Together, they form a network for processing environmental representations of spatial layout and the relational geometry of the environment. The cell types most strongly implicated by human cognitive GWAS are hippocampal CA1 pyramidal neurons, cortical pyramidal neurons, and striatal medium spiny neurons. These neurons perfectly map onto these brain regions. This further indicates that the genes identified by GWAS act through the same neural systems as those identified by behavioural and imaging studies.

2.5) Mouse Molecular Genetics of Learning and Memory

Mouse genetics currently provides the most detailed mechanistic understanding of cognition. It is a reference point against which every other species is compared. Based on the Morris Water Maze test of hippocampal-dependent spatial learning, Tsien (1996) performed a CA1-specific knockout of the NMDA receptor subunit GRIN1. This impaired spatial ability and eliminated NMDA receptor-dependent long-term potentiation. This did not affect non-spatial learning. This was one of the first direct demonstrations that a specific molecular machinery in a specific brain region is necessary for a specific cognitive function. Tang (1999), went in the opposite direction and created 'Doogie Mice' by overexpressing the GRIN2B gene in the forebrain. This resulted in improved novel-object recognition,

fear memory, and water-maze performance via prolonged NMDA receptor activation.

Bourtchuladze (1994) demonstrated the role of the CREB1 gene in long-term memory formation. CAMK2A was established as a core plasticity gene by Silva and his colleagues. Yamasaki (2009) found that mice lacking one copy of the CAMK2A gene show working memory deficits. Heldt (2007) demonstrated that the specific deletion of the BDNF gene in the adult hippocampus impairs spatial memory.

2.6) Other Species in the Study Panel

All the other species in the study panel showcase behaviour that requires high fluid reasoning and spatial ability. These include chimpanzees solving novel problems with tools (Hopkins, 2014), New Caledonian crows manufacturing hooked tools from twigs (Hunt, 1996; Taylor, 2009), numerical abilities in the Rhesus Macacque (Brannon & Terrace 1998), long distance spatial memory and mirror self recognition by African Elephants (Plotnik,2006), problem solving and explorative behaviour by the Great Tit (Kim, 2018), spatial navigation and associative learning by zebrafish (Best&Alderton 2008), human directed social cognition by dogs (vanHoldt, 2017) and fruit flies forming associative memories (Davis, 2005).

For chimpanzees and other primates, the only quantitative genetic analysis of intelligence is Hopkins (2014). They tested 99 chimpanzees across 13 cognitive tasks and estimated the heritability of the general 'g' factor to be 50%. Since no cognitive GWAS has been conducted in any other primate, the genetic architecture of cognition in our closest relatives is essentially unknown.

For canids, vonHoldt (2017) identified variations in the GTF2I and GTF2IRD1 genes- the canine version of the Williams-Beuren syndrome region. These appear to drive human-directed social behaviour in dogs while remaining unchanged in wolves. MacLean (2019) analysed over 14,000 dogs and identified 131 SNPs associated with personality differences that are involved in neurogenesis and hippocampal development. Gnanadesikan (2020) identified 188 genes explaining cognitive differences between breeds and found inhibitory control (the ability to resist impulses) to be 70% heritable.

For birds, Semenov (2024) combined experimental spatial cognition testing in wild mountain chickadees with the whole-genome sequencing of 162 birds. They identified 97 genes that are associated with spatial learning and memory. The Rho Family GTPases pathway, which is essential for neuronal development, dendritic arborisation and hippocampal neurogenesis, was overrepresented. Branch (2022) provided 15 replicated genes from a smaller cohort study of 42 birds. Together, these two studies represented the only published cognitive GWAS in a wild species.

For fruit flies, the learning and memory literature comes from Benzer's classical memory mutants. It has expanded to encompass multiple genes involved in the cAMP-PKA-CREB cascade, NMDA-related glutamate signalling, and mushroom body function (Davis, 2005; Heisenberg, 2003). Mushroom body function is to integrate sensory stimuli to facilitate associative learning, memory storage, and adaptive

decision-making in the invertebrate brain.

2.7) Impact of Current Work

The large-scale GWAS literature has established that cognitive ability is highly polygenic, with multiple genes contributing in small amounts. Although polygenic scores can only explain between 7-10 % of the variance in cognitive performance, they have become a standard research tool across psychology, psychiatry, and education research. This includes studies about the genetic overlap between intelligence and mental disorders (Lam, 2019). The identification of neurogenesis and synaptic plasticity as strongly enriched pathways has directed mechanistic research towards specific molecular targets and has contributed to various models of Alzheimer's disease. Mouse molecular genetics has provided a foundation for interpreting mechanistic and comparative work in other species and producing candidate targets for cognitive enhancement research. The chickadee GWAS has demonstrated the potential of large-scale GWAS in other wild species. Canid cognitive genetics has practical applications such as assistant dog selection (Gnanadesikan, 2023).

2.8) Gaps in Literature

One of the biggest gaps in modern human cognitive genetics is the missing heritability problem as GWAS only explains a fraction of the heritability of (g), estimated by twin studies. Another gap, called the fine-mapping problem, concerns whether GWAS findings actually identify causal genes or merely tag genomic regions that contain causal variants. This remains unsolved for intelligence so most of the 350 genes in this study's primary set should be taken as candidates rather than confirmed causal contributors.

Several other important gaps in the literature also motivate this study. No large-scale GWAS has targeted fluid reasoning as the primary phenotype, forcing the use of proxies. There is no well-powered human spatial-cognition GWAS, which is why the study uses a bidirectional approach. There is no large-scale cognitive GWAS in other primates. There is also an absence of systematic cross-species ortholog analyses at the gene level. The closest work to this is Hill (2016), which statistically demonstrated that SNPs associated with cognition cluster in evolutionarily constrained genomic regions, but this analysis operated at the SNP and large-region levels. This was not at the level of specific genes mapped to specific species.

Few published studies have systematically taken a human cognitive GWAS gene set and mapped both its ortholog conservation and its functional cognitive roles across a broad phylogenetic panel or conducted rigorous gene-level overlap analyses between human GWAS genes and independently compiled cognitive gene sets from other species. Another gap is the conservation-function interpretive problem. Many comparative studies treat the presence of an ortholog as evidence that it performs the same function in that species. This assumption has rarely been tested systematically. Another is the claim that cognitive pathways converge across species. This has rarely been tested against a formal statistical baseline, so it is possible that the observed overlap is due to chance.

3) METHODOLOGY

This study is a secondary data analysis. No primary data has been collected, and no experiments or surveys were conducted. Instead, the study applies new computational and statistical analyses to existing and compiled published datasets and genomic databases in order to answer the research question.

The first dataset comes from two large-scale GWAS of general cognitive function. Davies (2018) studied 300,486 participants from the UK Biobank. They identified 709 genes associated with cognitive ability using MAGMA (de Leeuw, 2015), a gene-based analysis tool that aggregates signals from all SNPs within a window around each gene to a single gene- level statistic. Savage (2018) studied 269,867 participants across multiple cohorts and identified 507 genes using this same method. For both studies the MAGMA gene-based lists were used, not the larger sets from Savage's positional, eQTL, and chromatin (FUMA) mapping (~1,016 genes), so the lists are method-matched before intersection. These two studies were selected because they are the largest published GWAS of cognitive function. This gives them the statistical power to detect genes with small effects on cognition.

The second dataset comprises ortholog annotations from the Ensembl Compara database (release 115). This is a curated database that identifies corresponding genes across species using algorithmic pipelines that compare protein sequences and genomic context. It was selected because it covers all major vertebrate and invertebrate model organisms and provides standardised ortholog calls.

The third source is the NCBI RefSeq genome assembly for the New Caledonian crow (assembly GCF_009650955.1). This species is not included in the Ensembl Compara pipeline; its proteome was downloaded from NCBI and used as the target for DIAMOND protein alignments.

The fourth dataset comes from two studies of spatial cognition in mountain chickadees (Poecile gambeli): Branch (2022; N = 42, 15 replicated genes) and Semenov (2024; N = 162, 97 genes with characterised functions). The 99-gene chickadee set used in this study was compiled by taking the union of the 97 Semenov and 15 replicated Branch genes after removing duplicates and genes without human orthologs. This enables the bidirectional analysis starting from genes associated with cognition in the chickadee and mapping back to the human cognitive gene set. This makes the study comparative rather than just a top-down conservation survey.

The fifth dataset comes from phenotype databases and published reviews of learning and memory in mice and fruit flies. These helped construct the cognitive gene sets for the cross-species comparison. For mice, the primary source was the comprehensive review by Lee (2014). This was supplemented by canonical Silva-laboratory studies, the MGI database and IMPC genomic resources. This helped produce a list of 56 genes with established roles in learning and memory. For fruit flies, the FlyBase database and established Drosophila neuroscience literature were used to compile the 49 learning and memory genes (mapping to 47 unique human orthologs). This includes work on dunce, rutabaga, and the cAMP–PKA–CREB

pathway. It was important that both these gene sets came entirely from literature independent of this study's own 350-gene list, to avoid circularity.

The sixth dataset consists of Gene Ontology (GO) enrichment data obtained using the g: Profiler tool (Raudvere 2019). The statistical tests described in the Model section were implemented in Python 3.11 using Fisher's exact test and the hypergeometric distribution. Additionally, 13 candidate genes for human spatial cognition were compiled from smaller-scale studies. This includes the hippocampal volume GWAS by Hibar (2017; N = 33,536), candidate gene studies of spatial navigation and the Williams syndrome deletion region (Hirota, 2003, and Dai, 2009). These have been included to compensate for the lack of a large-scale human spatial-cognition GWAS; all 13 are presented in Section 4.6. .

3.1) Analytical Model

The study uses a comparative ortholog conservation framework with three layers.

The first layer asks whether human cognitive genes are conserved in other species. If yes, then are they more conserved than genes in general? For each of the 350 human genes, it was checked whether an ortholog exists in each of the nine other species. The conservation rate of the cognitive gene set was then compared against that of all 23,262 human protein-coding genes (from Ensembl release 115) using one-sided Fisher's exact tests applied to a 2-by-2 contingency table. Three species were tested for this comparison. The Mouse (90 million years of divergence from humans), Zebrafish (450 million years), and fruit fly (600 million years). This unidirectional test was chosen as if cognitive genes are functionally important across evolution, then they should be more conserved than the genome average. The odds ratio quantifies the magnitude of enrichment; values greater than 1 indicate cognitive genes are more conserved than the genome average. The p-value indicates how likely the observed enrichment would be under the null hypothesis of no difference.

The second layer asks whether the genes associated with cognition in different species actually overlap. Cognitive gene sets were assembled for humans, mice, fruit flies, and mountain chickadees. For each of the six pairwise comparisons, hypergeometric distribution was used to test whether the observed overlap exceeds chance. The hypergeometric test calculates the probability of observing k or more shared genes between two sets of sizes m and n in a genome of N = 23,262 genes. For the human-mouse comparison (m = 350, n = 56 k = 11). The expected overlap under the null hypothesis of random sampling is (m × n) / N = (350 × 56) / 23,262 = 0.84 genes. This yields a fold enrichment of 11 / 0.84 ≈ 13.1, meaning the two sets share approximately 13 times more genes than chance would predict. For comparisons where zero genes overlapped — all three comparisons involving the chickadee — the depletion tail of the hypergeometric distribution was used, $P(X \leq 0)$, rather than the enrichment tail. This is because the enrichment test $P(X \geq 0)$ always equals 1.0 when zero genes overlap. The depletion test instead asks whether the absence of overlap is itself statistically significant. Because the overlap matrix involves six pairwise comparisons, the p-values were Bonferroni-corrected, yielding an adjusted significance threshold of $\alpha = 0.05/6 \approx 0.0083$. For the mouse-fly comparison, where the calculation underflowed in

double-precision floating-point arithmetic, the p-value was recomputed in log space using (scipy.stats.hypergeom.logsf). This yielded an upper bound of $p < 1 \times 10^{-16}$.

The third layer asks whether biological pathways are shared between species, even if the specific genes differ. GO enrichment analysis was run separately for the human, mouse, and fly cognitive gene sets, each against the whole-genome background, at a false discovery rate threshold of 0.05 (Benjamini-Hochberg correction). The GO terms significantly enriched across all three species were identified as convergent pathways. This convergence is formally tested with a permutation analysis (Section 4.5) where random human gene sets of matched size were enriched under the identical pipeline to build a null distribution for the three-way term overlap.

3.2) Sampling and Selection Criteria

The human gene set was made by the intersection of Davies (2018) and Savage (2018) GWAS studies, giving a total of 352 gene symbols. Three of these symbols (TMEM110, MUSTN1, and TMEM110-MUSTN1) mapped to the same Ensembl gene identifier (ENSG00000272573), giving 350 genes. The 511 genes that appeared in only one of the studies have been excluded to reduce the false-positive rate.

Because 19 of the 352 original symbols had been renamed since the source studies were published (for example, HARS became HARS1 and SEPT3 became SEPTIN3), all subsequent analyses were then matched on Ensembl identifiers rather than gene names to avoid undercounting due to unstable nomenclature.

A gene was counted as conserved if any ortholog relationship was present, regardless of whether it was one-to-one, one-to-many, or many-to-many. This is because the primary question is whether the gene exists in a given species at all, which is why the conservation rates are upper-bound estimates. All non-human gene sets in the overlap analysis were mapped to human orthologs prior to the comparison, placing everything in a common identifier space, yielding N = 23,262 protein-coding genes. For the GO enrichment, all three ontology domains were included, and a convergent term was defined as any term that reached significance (g: SCS-adjusted $p < 0.05$) in the human, mouse, and fly gene sets simultaneously.

Table 1 lists the species panel and the conservation results for each. Divergence times are drawn from the TimeTree database (Kumar 2017).

Table 1. Ortholog conservation of 350 human cognitive GWAS genes across nine species. Conservation means that at least one ortholog was identified. For eight species, this was determined via Ensembl Compara. For the New Caledonian crow, it was determined by Reciprocal Best Hit (RBH) DIAMOND alignment against the NCBI RefSeq proteome. The mean identity is the average target-to-source amino acid identity. The divergence times are approximate and sourced from TimeTree. The "Singletons"

column gives the number of one-to-one orthologs for Compara species and the number of RBH for the New Caledonian crow.

	Species	Common Name	Divergence (Myr)	Orthologs	Conservation (%)	Singletons	Mean Identity (%)	Method
1	Pan troglodytes	Chimpanzee	≈6	344	98.3	329	97.3	Compara
2	Macaca mulatta	Rhesus macaque	≈25	335	95.7	324	94.6	Compara
3	Mus musculus	Mouse	≈90	337	96.3	328	87.6	Compara
4	Canis lupus familiaris	Dog (wolf proxy)	≈95	332	94.9	317	88.0	Compara
5	Loxodonta africana	African elephant	≈100	326	93.1	314	84.4	Compara
6	Corvus moneduloides	NC Crow	≈320	301	86.0*	245	73.2	RBH-DIAMOND
7	Parus major	Great Tit	≈320	293	83.7	280	71.1	Compara
8	Danio rerio	Zebrafish	≈450	288	82.3	197	57.7	Compara
9	Drosophila melanogaster	Fruit Fly	≈600	199	56.9	66	31.9	Compara

Of the 301 crow orthologs, 245 come from RBH and 56 from forward-only hits; tested against the great tit Compara results, the RBH calls showed 88.4% accuracy, but the one-directional set lacks separate verification, so relying only on reciprocal hits, the conservation drops to 70.0% (245/350). Both figures are included for clarity. Notably, entries under Singletons (*) reflect only crow-side RBH numbers, not strict Compara-defined one-to-one relationships.

The conservation gradient requires attention to three points. The zebrafish set contains 288 orthologs, but 91 genes appear more than once due to an ancient whole-genome duplication event in teleost fish around 320 million years ago; comparing its 197 single-copy cases directly with those in mammals introduces bias. The domestic dog genome (Canis lupus familiaris) has been used instead of the grey wolf, as there is

no fully annotated wolf version available with Compara orthology mappings; the confidence remains high as both share nearly identical protein sequences of around 99.9% (Freedman, 2014). One entry, CCDC101 (ENSG00000131789), showed empty results during database retrieval, potentially because its ID has been retired, so it is not in the final conservation tally.

The four cognitive gene sets used in the overlap analysis are summarised in Table 2.

Table 2. Cross-species cognitive gene sets were used in the overlap analysis, where all non-human gene sets were mapped to human orthologs before comparison.

	Species	Genes (N)	Evidence Type	Primary Sources
1	Human	350	GWAS (common variants)	Davies et al.(2018); Savage et al.(2018)
2	Mouse	56	Knockout (loss-of-function)	Lee (2014); Silva laboratory; others
3	Fruit fly	47*	Classical Genetics	Drosophila learning/memory literature
4	Mountain chickadee	99	GWAS (common variants)	Branch et al. (2022); Semenov

*49 genes mapping to 47 unique human orthologs

3.3) Data Collection

Ortholog data for eight of the nine species were retrieved from Ensembl Compara via BioMart. For each of the 350 human genes (by Ensembl ID), BioMart returned all ortholog relationships in the target species. This includes the ortholog type, the target gene identifier, and the target-to-source amino acid identity. For the New Caledonian Crow, the orthologs were identified using DIAMOND v2.1.8 (Buchfink 2021) in very-sensitive mode (e-value 10^{-5}, query coverage $\geq$ 50%). The 350 human protein sequences were then aligned against the crow proteome (NCBI RefSeq GCF_009650955.1), and then reciprocally back against the human proteome. A gene pair was classified as an RBH ortholog if each sequence was the other's highest-scoring DIAMOND hit. The same procedure was applied to the great tit (Parus major) to validate the pipeline. The DIAMOND RBH calls matched Ensembl Compara calls at 88.4% precision. However, the recall was not calculated and remains a limitation. For the functional annotation, all 337 conserved mouse orthologs were queried against the MGI database and classified by manual curation of Mammalian Phenotype Ontology (MP) terms: a gene was classified as "cognitive" if its knockout produced phenotypes under MP:0002063 (abnormal learning/memory/conditioning) or any of its

descendant terms, and as "spatial" if the phenotype involved spatial learning paradigms such as the Morris water maze, Barnes maze, or radial arm maze.

4) FINDINGS AND DISCUSSION

This section develops the quantitative results presented in all 8 tables, further building on the paper's central argument. The findings map onto 5 related categories: the conservation gradient of human cognitive genes across nine species, the enrichment of cognitive genes relative to the genome background, the cross-species overlap between the compiled cognitive gene sets, the functional annotation of conserved orthologs in mouse and fruit fly, the spatial cognition candidate gene analysis. These results support the two-layer model of the genetic architecture of cognition in which a deeply conserved molecular core is shared across the animal kingdom, while natural selection for better cognition in different lineages can act on different gene sets.

4.1) Results

Table 3. Conservation enrichment of cognitive GWAS genes relative to all the human protein-coding genes. The background rates are from 23,262 genes in Ensembl. The odds ratios and p-values are from one-sided Fisher's exact tests.

	Common Name	Cognitive Conservation (%)	Background Conservation (%)	Odds Ratio	p-value
1	Mouse	96.3 (337/350)	76.2 (17717/23262)	8.25	1.07×10^{-25}
2	Zebrafish	82.3 (288/350)	60.6 (14096/23262)	3.06	8.88×10^{-19}
3	Fruit Fly	56.9 (199/350)	38.5 (8961/23262)	2.13	2.10×10^{-12}

Table 4. The Cross-species cognitive gene set overlap, significance tested using the hypergeometric distribution (N = 23,262). For zero-overlap cases, the depletion tail $P(X \leq 0)$ was used. The Bonferroni-corrected significance threshold for the six comparisons is $\alpha = 0.0083$.

	Comparison	Set Sizes	Shared	Expected	Fold	p-value	Significant

			Genes		Enrichment		(Bonferroni α=0.0083)
1	Human×Mouse	350 × 56	11	0.84	13.1	6.2×10^{-10}	Yes
2	Human×Fly	350 × 47	8	0.71	11.3	4.6×10^{-7}	Yes
3	Mouse×Fly	56 × 47	11	0.11	97.2	$<1\times10^{-16}$	Yes
4	Human×Chickadee	350 × 99	0	1.49	0	0.22	No
5	Mouse×Chickadee	56 × 99	0	0.24	0	0.79	No
6	Fly×Chickadee	47 × 99	0	0.20	0	0.82	No

Table 5. Functional annotation of conserved orthologs in the mouse and fruit fly. Spatial phenotype counts for mice are a subset of the cognitive phenotype count.

	Species	Conserved Orthologs	Cognitive Phenotype	Spatial Phenotype (subset)	Cognitive (%)
1	Mouse (MGI)	337	21	13	6.2
2	Fruit fly (FlyBase)	199	7	–	3.5

Table 6. Gene Ontology (GO) term enrichment and convergence across the three species with cognitive gene sets, counting GO terms significantly enriched at g:SCS-adjusted $p < 0.05$.

	GO Term Set	Significantly Enriched Terms
1	Human (350 gene set)	99

2	Mouse	725
3	Fruit Fly	214
4	Shared (human ∩ mouse)	88
5	Shared (human ∩ fruit fly)	56
6	Shared (mouse ∩ fruit fly)	152
7	Shared across all three	56

Table 7. The 13 human spatial-cognition candidate genes were examined against the analysed gene sets. "Appears in" indicates which of the analysed sets (the human, mountain chickadee, mouse, or fruit fly) contains the gene; "—" denotes none.

	Gene (alias)	Spatial-Cognition Role	Primary Evidence	Appears in	Strength
1	DPP4	Hippocampal Volume	Hibar, 2017	Human-350	Moderate
2	MAST4	Hippocampal Volume	Hibar, 2017	Human-350	Moderate
3	ASTN2	Hippocampal Volume (neuronal migration)	Hibar, 2017	-	Moderate
4	MSRB3	Hippocampal Volume	Hibar, 2017	-	Moderate
5	SHH (upstream)	Hippocampal Volume	Hibar, 2017	-	Moderate
6	GTF2IRD1	Visuspatial Construction (Williams syndrome)	Hirota, 2003; Dai, 2009	-	Strong (rare variant)
7	GTF2I	Williams region gene; normal-range effects	Hirota 2003; Dai, 2009	-	Strong (rare variant)
8	APOE	Navigation-strategy selection	Bohbot, 2016; Coughlan, 2019	Mouse	Strong
9	WWC1 (KIBRA)	Episodic memory, spatial navigation	Papassotiropoulos, 2006	-	Moderate

10	EPHX2	Spatial working memory	Zhang, 2022	-	Weak (unreplicated)
11	BDNF	Hippocampal plasticity, spatial memory	Candidate-gene studies	Mouse	Weak (unreplicated)
12	COMT	Working memory, prefrontal dopamine	Candidate-gene studies	Mouse	Weak (unreplicated)
13	NOS1	Nitric-oxide signalling in spatial memory	Weitzdoerfer, 2004	Chickadee	Moderate

Table 8. Permutation test of three-way pathway convergence. The observed overlap of 56 GO terms shared across the human, mouse, and fly cognitive gene sets was compared against a null distribution built from 290 random human gene sets of matched size (352 genes), enriched under an identical g:SCS-corrected pipeline, with the mouse and fly term sets held fixed. No random permutation reached the observed value.

	Permutation test statistic	Value
1	Observed three-way overlap (GO terms shared by human, mouse and fly gene sets)	56
2	Permutations (random human gene sets, 352 genes each)	290
3	Null distribution mean	1.65
4	Null distribution standard deviation	2.72
5	Null distribution maximum	16
6	Permutations reaching or exceeding observed value	0
7	Empirical p-value	< 0.0034
8	Departure from null expectation	≈ 20 SD

4.2) Human Cognitive Genes Are Significantly More Conserved Than the Genome Average

The conservation gradient in Table 1 tracks the expected phylogenetic pattern. This decreases as the evolutionary distance increases, from 98.3% in the chimpanzee to 56.9% in the fruit fly. This is not surprising, as any gene set should show less conservation as the divergence time increases. The interesting thing is shown in Table 3. This places the cognitive gene sets side by side with the genome-wide average for humans. In mice, 96.3% of cognitive genes are conserved, compared to just 76.2% for all genes. This yields an odds ratio of 8.25 and a significant p-value of 1.1×10^{-25}. The same pattern is followed by the zebrafish and fruit fly, where cognitive gene conservation exceeds the background conservation.

These enrichments are consistent across evolutionary distances spanning 90 to 600 million years, and highly statistically significant. This indicates that human cognitive GWAS genes are more likely to have orthologs in other species compared to random genes. This suggests they belong to an evolutionarily constrained subset of the genome. This extends the earlier finding by Hill (2016) that cognitive-associated SNPs cluster in conserved genomic regions, which was a SNP-level and region-level analysis, to the gene level with explicit per-species quantification. This is the first of the paper's empirical contributions.

This enrichment should be interpreted cautiously, and on its own it does not establish the "conserved toolkit" hypothesis. Elevated conservation is necessary but not sufficient evidence for that model, because genes with essential developmental, neurological, or housekeeping functions are also strongly conserved, and human cognitive genes are disproportionately brain-expressed, synaptically localised, and pleiotropically involved in core cellular processes. High conservation may therefore reflect purifying selection on those correlated properties rather than selection for cognition as such. The result is best read as consistent with, rather than proof of, the toolkit hypothesis. Distinguishing selection for cognitive function from selection for these correlated properties would require comparison against a control set matched for expression and constraint, which is identified here as a necessary next step.

4.3) Cross-Species Overlap: A Sharp Divide Between the Human-Mouse-Fly Triangle and the Chickadee

The main finding of this paper comes from the six-way pairwise overlap matrix. The results reveal two sharply distinct patterns.

The human, mouse, and fruit fly cognitive gene sets overlap at rates that exceed chance. The human-mouse comparison shows 11 shared genes, compared to an expected overlap of 0.84. The human-fly comparison shows 8 shared genes compared to an expected 0.71. The mouse-fly comparison shows 11 shared genes compared to an expected 0.11. All three of these p-values survive Bonferroni correction for the six comparisons at $\alpha = 0.0083$. The genes shared in these comparisons include PDE4D, NCAM1, SHANK3, GRIA4, MAPT, BAIAP2, DTNBP1, EP300, PLCB1, PTPRD, and SLC6A4, supporting the conserved cognitive toolkit model. Two of these genes, NCAM1 and PDE4D, appear in all

three pairwise overlaps of the human, mouse, and fly. They are the strongest single candidates in the entire analysis for conserved cognitive function across mammals and invertebrates. NCAM1 is a synaptic adhesion protein with established roles in long-term potentiation and memory formation. PDE4D hydrolyses cAMP as part of the cAMP-PKA-CREB signalling cascade that Kandel's work established as the molecular substrate of long-term memory. The mouse-fly overlap of 97.2-fold enrichment is significant because, despite them being separated by approximately 600 million years of evolution, their shared learning and memory genes include the core cAMP-PKA-CREB pathway components. These include CREB1, PDE4D, GSK3B, NCAM1, NF1, and CAMK2A, along with the NMDA receptor subunits GRIN1 and GRIN2B. This is quantitative evidence for Kandel's observation that the molecular machinery of long-term memory is conserved from invertebrates to mammals, now also extended with a null-model test.

There is one important caveat before interpreting the overlap matrix. The four gene sets are not derived by the same methods as human and chickadee gene sets come from GWAS whereas mouse and fly gene sets come from single gene knockouts of large effect mutations. These two approaches sample different parts of the effect-size and allele-frequency spectrum, so raw overlap counts cannot be compared directly across species pairs — a low overlap may simply reflect this methodological mismatch rather than true biological divergence. This heterogeneity makes the human-mouse-fly signal conservative rather than inflated as the three sets overlap far beyond chance despite being built by different methods, strengthening the evidence of a shared molecular core. However, the same heterogeneity also means that for the chickadee comparisons, a low or null overlap must be interpreted cautiously as it may reflect the GWAS/knockout mismatch.

The mountain chickadee spatial cognition gene set doesn’t share any genes with the other three species. The depletion p-values for these zero overlaps are 0.22, 0.79, and 0.82. None of these values are statistically significant. The chickadee’s lack of overlap is suggestive but cannot be definitively distinguished from chance using standard thresholds. The expected overlap of the chickadee gene set with that of the mouse and fly is small (0.24 with mouse and 0.20 with fly). Therefore, finding zero is not very significant. However, the chickadee-human comparison stands out due to an expected overlap of 1.49 genes and a depletion p-value of 0.22.

This creates an interesting dilemma and interpretive constraint. Despite the low p-value, the zero-overlap finding cannot be used to claim that chickadee spatial cognition genes are definitively distinct from the human, mouse, and fly cognitive gene sets. This is because it still does not meet the p<=0.05 threshold. It can only be used to say that there is no positive evidence of overlap. This is consistent with but does not prove the hypothesis that natural selection for spatial cognition in chickadees has acted on a distinct gene set.

Nevertheless, the contrast between the tight human-mouse-fly triangle and the lack of chickadee overlaps is still interesting. One possibility is that it reflects the different methodologies used to identify the gene sets. This is because the human and chickadee gene sets both come from GWAS of common variants in

natural populations, while the mouse and fly gene sets come from knockout studies of large-effect mutations. If overlap was only driven by methodological similarity, the human-chickadee comparison should have been the strongest overlap as both datasets are GWAS based. However, it is the weakest. Therefore, the most likely explanation is that the genes associated with natural spatial cognition variation in the chickadee are distinct from those associated with general cognitive variation in humans because the selection pressures acting on them are different (food-hoarding survival versus the polygenic pressures shaping human g). The second possibility is that the chickadee GWAS sample size of 162 birds is simply too small, and a larger study would find more genes and overlap.

4.4) Conservation of Sequence Does Not Imply Conservation of Function

Table 5 presents the functional annotation results. Of the 337 mouse orthologs of the 350 human cognitive genes, only 21 produce learning or memory deficits when knocked out (6.2%). Of those 21, 13 specifically affect spatial learning. In the fruit fly, only 7 of the 199 conserved orthologs (3.5%) have documented cognitive or neural roles in the established Drosophila literature. The remaining 94% of mouse orthologs and 96.5% of fly orthologs seem to have no known cognitive function. They instead serve a wide range of unrelated cellular processes. Eg- metabolism, cell cycle regulation, structural maintenance, etc.

This is important as it demonstrates that the conservation gradient in Table 1 cannot be read as evidence that human cognitive genes "are" cognitive genes in other species. The majority of conserved orthologs have essential non-cognitive functions that impose their own purifying selection, and they also happen to be cognitive in humans. This is because the machinery of cognition draws on many core cellular components. Any comparative genomics study that fails to make this distinction risks overinterpreting its findings. This is the "conservation-function gap" mentioned in the literature review section that this study addresses by directly combining conservation data with functional annotation.

There are two alternative explanations for the low functional annotation results as they set a limit of how strongly this gap can be interpreted. . Firstly, the figures of 6.2% for mouse and 3.5% for fly are lower bounds, not true rates as an ortholog was counted as cognitive if a knockout has actually been made and tested for a learning or memory phenotype, and most orthologs have never been examined this way. Therefore, a low percentage partly reflects the incomplete testing, not the confirmed absence of cognitive function. Secondly, genetic redundancy can mask cognitive roles as paralogues or parallel pathways may compensate for a single knockout and hide a phenotype that the gene still contributes to. The finding here is the sharp contrast between near-universal sequence conservation and the small documented-function fraction as many of the genes haven’t been rigorously tested so the paper does not claim that 94% and 96.5% of conserved orthologs are non-cognitive.

This finding has a further implication for the interpretation of the overlap matrix. The 11 human-mouse shared cognitive genes passed both the top-down human GWAS and in the independent bottom-up mouse knockout filter, which is why the overlap is 13.1-fold enriched ($p = 6.2 \times 10^{-10}$). These genes (PDE4D,

NCAM1, SHANK3, GRIA4, MAPT) are strong candidates for genes with conserved cognitive function across mammals, as they survived both methodological filters.

4.5) Pathway-Level Convergence Is Substantial and Exceeds Chance

Under a multiple-testing-corrected pipeline (GO:BP, GO:CC, GO:MF; g:SCS, $p < 0.05$), the human, mouse, and fruit fly cognitive sets were significantly enriched for 99, 725, and 214 GO terms respectively, of which 56 were shared across all three (Table 6). These convergent terms are heavily concentrated in synapse organisation and neuron development, including the postsynaptic density, regulation of trans-synaptic signalling, modulation of chemical synaptic transmission, neurogenesis, and neuron projection development. The pathway-level overlap far exceeds the gene-level overlap: where humans and mice share only 11 genes they share 88 enriched pathways, and where mouse and fly share 11 genes they share 152 pathways. This confirms that although the three species share few specific genes, they draw their cognitive function from broadly the same biological processes.

To test whether this convergence exceeds what large gene sets would share by chance, a permutation test was run using the same enrichment pipeline applied identically to every gene set. A null distribution was generated by drawing 290 random human gene sets of matched size (352 genes) from the 19,409 protein-coding genes in the genome, enriching each identically, and counting how many of their terms fell within the fixed observed mouse and fly term sets — a deliberately conservative design, since holding the large mouse and fly sets fixed gives each random set the maximum chance to overlap them. None of the 290 random sets reached the observed overlap (null mean 1.7 terms, maximum 16; Table 8), giving an empirical $p < 0.0035$. The observed convergence therefore lies roughly twenty standard deviations above the chance expectation and cannot be explained by the size of the gene sets alone. The pathways driving it are the same ones that have dominated the molecular learning-and-memory literature for decades: glutamate receptor signalling, cAMP cascades, synaptic plasticity, and neurodevelopment.

4.6) Spatial Cognition Candidates: A Single Point of Convergence with Chickadee

The analysis of 13 human spatial cognition candidate genes presented in Table 7, drawn primarily from hippocampal volume GWAS and the Williams syndrome literature, has yielded three findings. Firstly, two of the candidates, the DPP4 and MAST4 genes, already appear in the 350-gene human GWAS set. Both were originally identified as hippocampal volume loci by Hibar (2017). This confirms that at least some genes contributing to hippocampal structural variation also contribute to general cognitive function.

One candidate gene, NOS1, appears in the chickadee gene set. This is the only point of molecular convergence between the human spatial candidates and the chickadee gene set identified in the analysis. NOS1 is potentially a cognitive gene, as NOS1 knockout mice do show impaired spatial learning (Weitzdoerfer, 2004). However, a single gene is not enough to support a strong convergence claim, and NOS1 is not actually in the 350-gene set, so this point links a smaller candidate list to the chickadee rather than the primary gene set.

Three other candidate genes, APOE, BDNF, and COMT, appear in the mouse cognitive gene set but not in the primary human gene set; these are established genes for learning and memory in model organisms, but they do not show up in human cognitive GWAS because they are under strong purifying selection in humans. This is internally consistent with the ascertainment-bias interpretation. It provides indirect evidence that the missing heritability problem is partly driven by this kind of selection.

4.7) Limitations

There are eight limitations of this methodology that should be noted.

Firstly, the human source studies measured general cognitive function (g), rather than fluid reasoning ability. While the genetic correlation between the two is very high ($r_g \approx 0.88–1.0$), the gene set may also include variants associated with other cognitive domains. For spatial cognition, the concern is greater as the genetic correlation with g is lower (≈ 0.6) and part of the genetic variance in spatial ability is independent of g (Rimfeld, 2017), so a g-based set will miss spatial-specific genes — one reason the chickadee spatial-cognition data are analysed bidirectionally.

Secondly, the cross-species comparison brings fundamentally different types of genetic evidence. The human and chickadee genes come from GWAS of common variants. The mouse and fly genes, however, come from knockout studies. These approaches probe different segments of the allele-frequency and effect-size spectrum. The overlap between them is informative but not expected to be complete.

Thirdly, the chickadee GWAS specifically measures cache recovery accuracy (how well a bird remembers where it stored food). This is a specific spatial behaviour rather than a general spatial cognition trait. The sample size ($N = 162$) is also very small by human GWAS standards, and the 92% phenotypic variance explained raises the possibility of overfitting. However, the replication of 15 genes across two independent cohorts (Branch 2022; Semenov 2024) gives this some validation. The zero overlap between the chickadee and all other gene sets did not reach statistical significance ($p = 0.22$ for chickadee–human, hypergeometric depletion test). This indicates a result that is suggestive but not statistically distinguishable from chance.

Fourthly, several genes that are well known to be critical for learning and memory, such as CREB1, GRIN2B, and CAMK2A, appear in the mouse and fly gene sets but are absent from the human 350-gene set. This is likely because these genes are under strong purifying selection in humans, limiting the natural variation GWAS needs to detect them. This hypothesis is supported by the high loss-of-function intolerance scores ($pLI > 0.9$) for all three genes in the gnomAD constraint database (Lek 2016). This indicates a strong evolutionary constraint against protein-truncating variation. The absence of these genes from the human gene set does not indicate that they are unimportant for human cognition. Instead, it means GWAS cannot detect them because there is not enough common variation to test them.

Fifth, although the DIAMOND-based ortholog identification for the crow has been validated at 88.4% accuracy, it may include some false positive calls for the remaining fraction. The forward-only hit fraction (56 of 301) was not independently validated. Both these considerations mean that the crow conservation figure should be interpreted with caution.

Sixth, sequence conservation does not mean functional conservation: only 6.2% of mouse and 3.5% of fly orthologs have documented cognitive roles, and these are lower bounds set by incomplete annotation (Section 4.4).

Seventh, although the dog/wolf substitution is okay for protein-level conservation but cannot capture regulatory differences between wild and domestic canids that may affect gene expression.

Eighth, the hypergeometric overlap test assumes that the gene sets are random samples from the genome. However, the cognitive genes are systematically enriched for brain-expressed, synaptically localised, and evolutionarily constrained loci. This assumption is common in comparative genomics, and the qualitative contrast between significant overlap and non-significant overlap is robust to any uniform bias in p-value calibration. The Layer 3 pathway convergence was additionally validated with a permutation test (Section 4.5), which confirmed that the three-way overlap significantly exceeds chance under a consistent, multiple-testing-corrected enrichment pipeline.

5) CONCLUSION

This paper set out to answer whether the genes associated with fluid reasoning and spatial cognition in humans are conserved in other species that exhibit these cognitive abilities, and what the pattern of conservation reveals about the genetic architecture of intelligence across species. The evidence shows that human cognitive GWAS genes are conserved across species at rates that are much higher than the genome-wide background. The human, mouse, and fruit fly cognitive gene sets overlap at rates that significantly exceed chance and remain significant after Bonferroni correction. Contrastingly, the mountain chickadee shares no genes with any of the three other species. However, this zero-overlap result does not reach statistical significance and should be interpreted as suggestive rather than conclusive. Taken together, and with the caveat that elevated conservation is consistent with rather than proof of the model, this evidence supports the conserved toolkit view.

The paper has also shown that sequence conservation does not imply functional conservation, as most conserved cognitive genes appear to have non-cognitive functions in other species. However, at the pathway level, the convergence is much broader. This suggests that even where specific genes differ, the underlying biological processes can be shared. These findings contribute to the growing body of research suggesting that cognitive evolution proceeds primarily through regulatory and lineage-specific changes rather than through the invention of new cognitive genes. This paper attempts to extend this picture from

the pathway level to the gene level with explicit statistical testing that has rarely been performed in the comparative literature.

There are several implications of the research that are applicable to multiple fields. In cognitive neuroscience, the gene sets that overlap between human GWAS and mouse knockout studies provide a list of molecular targets for mechanistic research. For medical genetics, the absence of classic learning genes such as CREB1, GRIN2B, and CAMK2A from human GWAS gene lists suggests that rare-variant analyses will be essential to complete the picture of human cognitive genetic architecture and several of the human-mouse overlap genes are already implicated in human conditions (SHANK3 in autism, MAPT in tauopathies, NCAM1 in schizophrenia, and PDE4D as a target for memory enhancement), reinforcing their value for therapeutic research. A permutation test reported here (Section 4.5) confirms that the pathway-level convergence significantly exceeds chance; a remaining analysis that would further strengthen the conservation claim is a control comparing cognitive-gene conservation against genes matched for expression and constraint. Funding agencies should prioritise cognitive GWAS in other species. The chickadee work by Branch (2022) and Semenov (2024) has demonstrated that such studies are feasible with sample sizes in the low hundreds, and a larger cognitive GWAS in mountain chickadees or in a related food-caching species ($N > 500$) would help resolve whether the zero-overlap finding reflects real divergence. A cognitive GWAS in any primate, building on Hopkins(2014) demonstration that chimpanzee cognitive ability is approximately 50% heritable, would help test whether the human-mouse-fly overlap triangle extends to our closest evolutionary relatives. Finally, standardising how cognitive gene lists and phenotype annotations link to ortholog databases would make comparative work of this kind substantially easier to reproduce and extend.

REFERENCES

[1] Best, J. D., & Alderton, W. K. (2008). Zebrafish: an in vivo model for the study of neurological diseases. *Neuropsychiatric Disease and Treatment*, 4(3), 567–576.

[2] Bouchard, T. J., & McGue, M. (1981). Familial studies of intelligence: a review. *Science*, 212(4498), 1055–1059.

[3] Bourtchuladze, R., et al. (1994). Deficient long-term memory in mice with a targeted mutation of the cAMP-responsive element-binding protein. *Cell*, 79(1), 59–68.

[4] Branch, C. L., et al. (2022). The genetic basis of spatial cognitive variation in a food-caching bird. *Current Biology*, 32(1), 210–219.

[5] Brannon, E. M., & Terrace, H. S. (1998). Ordering of the numerosities 1 to 9 by monkeys. *Science*, 282(5389), 746–749.

[6] Buchfink, B., Reuter, K., & Drost, H. G. (2021). Sensitive protein alignments at the tree-of-life scale using DIAMOND. *Nature Methods*, 18(4), 366–368.

[7] Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge University Press.

[8] Cattell, R. B. (1963). Theory of fluid and crystallised intelligence. *Journal of Educational Psychology*, 54(1), 1–22.

[9] Chabris, C. F., et al. (2012). Most reported genetic associations with general intelligence are probably false positives. *Psychological Science*, 23(11), 1314–1323.

[10] Charrier, C., et al. (2012). Inhibition of SRGAP2 function by its human-specific paralogs induces neoteny during spine maturation. *Cell*, 149(4), 923–935.

[11] Cui, Y., et al. (2023). Cellular evolution of the human brain. *Neuron*, 111(22), 3560–3580.

[12] Dai, L., et al. (2009). Is it Williams syndrome? *GTF2IRD1* implicated in visual-spatial construction and *GTF2I* in sociability revealed by high resolution arrays. *American Journal of Medical Genetics Part A*, 149A(3), 302–314.

[13] Davies, G., et al. (2018). Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. *Nature Communications*, 9(1), 2098.

[14] Davis, R. L. (2005). Olfactory memory formation in *Drosophila*. *Annual Review of Neuroscience*, 28, 275–302.

[15] De Leeuw, (2015). MAGMA: Generalised gene-set analysis of GWAS data. *PLoS Computational Biology*, 11(4), e1004219.

[16] Dussex, N., et al (2021). Population genomics of the critically endangered kākāpō. Cell Genomics, 1(1), 100002.

[17] Enard, W., et al. (2002). Molecular evolution of FOXP2, a gene involved in speech and language. *Nature*, 418(6900), 869–872.

[18] Florio, M., et al. (2015). Human-specific gene ARHGAP11B promotes basal progenitor amplification and neocortex expansion. *Science*, 347(6229), 1465–1470.

[19] Freedman, A. H., et al. (2014). Genome sequencing highlights the dynamic early history of dogs.

PLoS Genetics, 10(1), e1004016.

[20] Gnanadesikan, G. E., et al. (2020). Estimating the heritability of cognitive traits across dog breeds reveals highly heritable inhibitory control and communication factors. *Animal Cognition*, 23(5), 953–964.

[21] Gnanadesikan, G. E., et al. (2023). Breed differences in dog cognition associated with brain-expressed genes and neurological functions. *Integrative and Comparative Biology*, 60(4), 976–990.

[22] Heisenberg, M., et al.(2003). Mushroom body memoir: from maps to models. *Nature Reviews Neuroscience*, 4(4), 266–275.

[23] Heldt, S. A. (2007). Hippocampus-specific deletion of BDNF in adult mice impairs spatial memory and extinction of aversive memories. *Molecular Psychiatry*, 12(7), 656–670. [VERIFY]

[24] Hibar, D. P., et al. (2017). Novel genetic loci associated with hippocampal volume. *Nature Communications*, 8, 13624.

[25] Hill, W. D., et al. (2016). Molecular genetic aetiology of general cognitive function is enriched in evolutionarily conserved regions. *Translational Psychiatry*, 6(12), e980.

[26] Hill, W. D., et al. (2019). A combined analysis of genetically correlated traits identifies 187 loci and a role for neurogenesis and myelination in intelligence. *Molecular Psychiatry*, 24(2), 169–181.

[27] Hirota, H., (2003). Williams syndrome deficits in visual spatial processing linked to *GTF2IRD1* and *GTF2I* on chromosome 7q11.23. *Genetics in Medicine*, 5(4), 311–321.

[28] Hopkins, W. D., Russell, J. L., & Schaeffer, J. (2014). Chimpanzee intelligence is heritable. *Current Biology*, 24(14), 1649–1652

[29] Hunt, G. R. (1996). Manufacture and use of hook-tools by New Caledonian crows. *Nature*, 379(6562), 249–251.

[30] Jung, R. E., & Haier, R. J. (2007). The Parieto-Frontal Integration Theory (P-FIT) of intelligence. *Behavioral and Brain Sciences*, 30(2), 135–154.

[31] Kandel, E. R. (2001). The molecular biology of memory storage. *Science*, 294(5544), 1030–1038.

[32] Kim, J. M., Santure (2018). A high-density SNP chip for genotyping great tit (*Parus major*) populations and its application to studying the genetic architecture of exploration behaviour. *Molecular Ecology Resources*, 18(4), 877–891.

[33] King, M. J., et al. (2019). Genetic and environmental influences on spatial reasoning: A meta-analysis of twin studies. *Intelligence*, 73, 65–77.

[34] Konopka, G., et al. (2009). Human-specific transcriptional regulation of CNS development genes by FOXP2. *Nature*, 462(7270), 213–217.

[35] Kumar, S., (2017). TimeTree: A resource for timelines, timetrees, and divergence times. *Molecular Biology and Evolution*, 34(7), 1812–1819.

[36] Lam, M., et al. (2019). Comparative genetic architectures of schizophrenia in East Asian and European populations. *Nature Genetics*, 51(12), 1670–1678.

[37] Lee, J. J., et al. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics*, 50(8), 1112–1121.

[38] Lee, Y. S. (2014). Genes and signalling pathways involved in memory enhancement in mutant mice. *Molecular Brain*, 7, 43.

[39] Lek, M., et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, 536(7616), 285–291.

[40] Liu, J.,(2023). Meta-analysis of genome-wide association studies of hippocampal volume uncovers novel loci and reveals new insights into brain aging. *Genome Medicine*, 15(1), 47.

[41] MacLean, E. L., et al. (2019). Highly heritable and functionally relevant breed differences in dog behaviour. *Proceedings of the Royal Society B*, 286(1912), 20190716.

[42] Montgomery, S. H., (2011). Adaptive evolution of four microcephaly genes and the evolution of brain size in anthropoid primates. *Molecular Biology and Evolution*, 28(1), 625–638.

[43] Morris, R. G. M. (1981). Spatial localisation does not require the presence of local cues. *Learning and Motivation*, 12(2), 239–260.

[44] Nithianantharajah, J., et al. (2013). Synaptic scaffold evolution generated components of vertebrate cognitive complexity. *Nature Neuroscience*, 16(1), 16–24.

[45] Okbay, A., (2022). Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nature Genetics*, 54(4), 437–449.

[46] Pfenning, A. R., (2014). Convergent transcriptional specialisations in the brains of humans and song-learning birds. *Science*, 346(6215), 1256846.

[47] Plomin, R., & Deary, I. J. (2015). Genetics and intelligence differences: Five special findings. *Molecular Psychiatry*, 20(1), 98–108.

[48] Plotnik, J. M.,(2006). Self-recognition in an Asian elephant. *Proceedings of the National Academy of Sciences*, 103(45), 17053–17057.

[49] Raudvere, U., (2019). g:Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Research*, 47(W1), W191–W198.

[50] Rimfeld, K., et al. (2017). Phenotypic and genetic evidence for a unifactorial structure of spatial abilities. *Proceedings of the National Academy of Sciences*, 114(11), 2777–2782.

[51] Savage, J. E., et al. (2018). Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nature Genetics*, 50(7), 912–919.

[52] Semenov, G. A., et al. (2024). The genetic basis of spatial learning and memory in a wild bird. *Current Biology*.

[53] Silva, A. J., Paylor, R., Wehner, J. M., & Tonegawa, S. (1992). Impaired spatial learning in alpha-calcium-calmodulin kinase II mutant mice. *Science*, 257(5067), 206–211.

[54] Sniekers, S., et al. (2017). Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. *Nature Genetics*, 49(7), 1107–1112.

[55] Spearman, C. (1904). "General intelligence", objectively determined and measured. *The American Journal of Psychology*, 15(2), 201–292.

[56] Tang, Y. P., et al. (1999). Genetic enhancement of learning and memory in mice. *Nature*, 401(6748), 63–69.

[57] Taylor, A. H., (2009). Complex cognition and behavioural innovation in New Caledonian crows. *Proceedings of the Royal Society B: Biological Sciences*, 277(1694), 2637–2643.

[58] Tsien, J. Z., et al. (1996). The essential role of hippocampal CA1 NMDA receptor-dependent synaptic plasticity in spatial memory. *Cell*, 87(7), 1327–1338.

[59] vonHoldt, B. M., et al. (2017). Structural variants in genes associated with human Williams-Beuren syndrome underlie stereotypical hypersociability in domestic dogs. *Science Advances*, 3(7), e1700398.

[60] Wainschtein, P., et al. (2022). Assessing the contribution of rare variants to complex trait heritability from whole-genome sequence data. *Nature Genetics*, 54(3), 263–273.

[61] Weitzdoerfer, R., (2004). Neuronal nitric oxide synthase knock-out mice show impaired cognitive performance. *Nitric Oxide*, 10(3), 130–140.

[62] Yamasaki, N., Maekawa, (2009). Alpha-CaMKII deficiency causes immature dentate gyrus, a novel candidate endophenotype of psychiatric disorders. *Molecular Brain*, 2, 6.

[63] Zhang, W., (2022). Genome-wide association study identifies a novel locus associated with spatial working memory. *Frontiers in Genetics*, 13, 893359.

[64] Zhao, B., (2019). Genome-wide association analysis of 19,629 individuals identifies variants influencing regional brain volumes and refines their genetic co-architecture with cognitive and mental health traits. *Nature Genetics*, 51(11), 1637–1644.