Polyglucosan body density in the aged mouse hippocampus is controlled by a novel modifier locus on chromosome 1

A. Gómez-Pascual[1]*, D. M. Glikman[2,3]*, H. X. Ng[4]*, J. E. Tomkins[5,6], L. Lu[7], Y. Xu[8,9], D. G. Ashbrook[7], C. Kaczorowski[10], G. Kempermann[11], J. Killmar[7], K. Mozhui[7,12], R. Aebersold[13], E. G. Williams[14], R. W. Williams[7#], R. W. Overall[15], M. Jucker[8,9], D. E. M. de Bakker[16#]

[1] Department of Information and Communications Engineering, University of Murcia, Murcia, Spain
[2] University of Vienna, Vienna, Austria
[3] Department of Neuronal Cell Biology, Center for Brain Research, Medical University of Vienna, Vienna, Austria
[4] Department of Cognitive Science University of California, San Diego, USA
[5] Centre for Misfolding Diseases, Department of Chemistry, University of Cambridge, Cambridge, UK
[6] Aligning Science Across Parkinson's (ASAP) Collaborative Research Network, Chevy Chase, MD, 20815
[7] Department of Genetics, Genomics and Informatics, University of Tennessee Health Center, Memphis, TN, USA
[8] Department of Cellular Neurology, Hertie-Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany
[9] German Center for Neurodegenerative Diseases (DZNE), Tübingen, Germany
[10] University of Michigan, Michigan, USA
[11] German Center for Neurodegenerative Diseases (DZNE), Dresden, Germany
[12] Department of Preventive Medicine, College of Medicine, University of Tennessee Health Center, Memphis, TN, USA
[13] Department of Biology, Institute of Molecular Systems Biology, ETH Zurich. Zurich, Switzerland
[14] Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Belval, Luxembourg
[15] Humboldt University of Berlin, Berlin, Germany
[16] Leibniz Institute on Aging – Fritz Lipmann Institute (FLI), Jena, Germany
* these authors contributed equally
[#] Corresponding authors

Abstract

Aging can be associated with the accumulation of hypobranched glycogen molecules (polyglucosan bodies, PGBs), particularly in astrocytes of the hippocampus. While PGBs have a detrimental effect on cognition in diseases such as adult polyglucosan body disease and Lafora disease, the underlying mechanism and clinical relevance of age-related PGB accumulation remains unknown. Here, we have investigated the genetic basis and functional impact of age-related PGB accumulation in 32 fully sequenced BXD-type strains of mice which exhibit a 400-fold variation in PGB burden in 16-18 month old females. We mapped a major locus controlling PGB density in the hippocampus to chromosome 1 at 72–75 Mb (linkage of 4.9 –logP), which we defined as the *Pgb1* locus. To identify potentially causal gene variants within *Pgb1*, we generated extensive hippocampal transcriptome datasets and identified two strong candidate genes for which mRNA correlates with PGB density—*Smarcal1* and *Usp37*. In addition, both *Smarcal1* and *Usp37* contain non-synonymous allele variations likely to impact protein function. A phenome-wide association analysis highlighted a trans-regulatory effect of the *Pgb1* locus on expression of *Hp1bp3*, a gene known to play a role in age-related changes in learning and memory. To investigate the potential impact of PGBs on cognition, we performed conditioned fear memory testing on strains displaying varying degrees of PGB burden, and a phenome-wide association scan of ~12,000 traits. Importantly, we did not find any evidence suggesting a negative impact of PGB burden on cognitive capacity. Taken together, we have identified a major modifier locus controlling PGB burden in the hippocampus and shed light on the genetic architecture and clinical relevance of this strikingly heterogeneous hippocampal phenotype.

Introduction

Polyglucosan bodies (PGBs) are formed by the aggregation of hypobranched glycogen molecules. Accumulation of PGBs is associated with neurodegenerative diseases including adult polyglucosan body disease (APBD) and Lafora disease (Duran et al., 2014). APBD is a rare neurogenetic disorder in which mutations in glycogen branching enzyme 1 (*GBE1*) reduce the efficiency of glycogen molecule synthesis, leading to hypobranched glycogen molecules prone to aggregation (Nitschke et al., 2022; Zebhauser et al., 2022). These aberrant glycogen chains aggregate in neurons and astrocytes, leading to cellular dysfunction across the peripheral and central nervous systems (Duran and Guinovart, 2015). As a result, APBD patients suffer severe peripheral neuropathy, muscle weakness and dementia. In Lafora disease, caused by mutations in epilepsy progressive myoclonus type 2a (*EPM2A)* and epilepsy progressive myoclonus type 2b (*EPM2B)* genes, PGBs accumulate in many organs and tissues including the brain, skin, liver, heart, and skeletal muscle. Typically, this leads to ataxia and seizures in adolescence and the progressive development of severe dementia (Duran and Guinovart, 2015; Moreno-Estellés et al., 2023; Wang et al., 2007).

Besides playing a role in neurological diseases, PGBs arise throughout the aging brain under non-pathological conditions (Duran and Guinovart, 2015). These PGBs are most prominent in astrocytes of the hippocampus, but can also be found in other regions and cells of the periventricular and subpial regions of the aged human brain (Duran and Guinovart, 2015). The astrocyte-specific accumulation of PGBs is likely due to the central role these glial cells play in glycogen homeostasis in the brain (Palmer and Ousman, 2018).

Astrocytes are a heterogeneous cell population with distinct functions, which we are only now starting to unravel (Viana et al., 2023). For example, a subset of hippocampal astrocytes have recently been identified that show an age-related autophagy dysregulation, leading to severely inflated autophagosomes. Although a direct relevance to cognitive function was not shown, the astrocytic PGB-like inclusion bodies had detrimental effects on the number of synapses in adjacent neurons (Lee et al., 2022). Considering that hippocampal astrocytes play an active role in memory consolidation and the location-specific encoding of the reward system (Doron et al., 2022), it seems likely that astrocytic dysfunctions play a key role during age-related cognitive decline. Therefore, the accumulation of PGBs in astrocytes of the hippocampus could have far reaching consequences on brain health and cognition.

Although PGBs have been studied extensively within the context of APBD and Lafora disease (Duran, 2023), there remains a lack of comprehensive understanding regarding the genetic architecture and pathophysiological significance of PGBs during the aging process. Here, we investigate the genetic architecture underlying PGBs by analysing the age-related occurrence of PGBs in 32 different BXD recombinant inbred mouse strains. One of the primary uses of the BXD family is to map genetic regions, also called quantitative trait loci (QTL), which influence phenotypic variation of a complex trait (Peirce et al., 2004; Taylor, 1978). Here, we employ QTL mapping to investigate the genetic architecture and pathophysiological significance of PGB burden, by exploiting novel transcriptomic and proteomic datasets on the aged BXD hippocampus as well as BXD behavioral data (Ashbrook et al., 2021).

Results

Variation in hippocampal PGB density maps to a gene-rich locus on chromosome 1

To investigate the variation of PGB density in the aged mouse brain, we quantified the number of PGBs in the hippocampus of 142 aged females (16–18 months old) sampled from 32 BXD strains and both parental strains C57BL/6J (B6) and DBA/2J (D2). The data can be found using genenetwork.org (GN) trait ID: BXD_10685. The average number of PGBs was determined from an average of four animals per strain (**Figure S1A**). Average PGB counts varied over 400-fold among BXD strains, with a variation between 1.5 and 646 PGBs per hippocampus (**Figure 1A-C**). Extraordinarily high PGB burden was observed in B6 and BXD20 strains, while PGBs were nearly undetectable in D2 and BXD32 strains (**Figure 1D-F**). Next, we asked which genetic loci contribute to differences in PGB density. To this end, PGB burden was mapped using sequence-based markers and a linear mixed model regression (GEMMA) which corrects for kinship (Ashbrook et al., 2021; Zhou and Stephens, 2012). The PGB numbers among strains is strongly skewed and values were therefore $\log_{10}$ transformed prior to mapping (GN trait ID: BXD_10686). The highest association (–logP 4.9) was found on chromosome 1 (Chr 1) at 72.5 Mb with a 1.5 LOD confidence interval, ranging from 72.0 to 75.0 Mb (**Figure 2A-D**). Approximately 50% of the variation in PGB burden was explained by this locus (r = –0.51, p = 2.2 x 10^{-3}, n = 34 strains), hereafter referred to as the *Pgb1* locus (**Figure 2E**). The *Pgb1* locus includes 94 potential coding regions of which 47 encoded validated transcripts (42 protein coding, four microRNAs, and one long non-coding RNA) (**Figure 3A**). *Table S1* lists the key data on each of the 47 candidates including mRNA and protein expression in the BXD hippocampus as well as the location of potential non-synonymous SNPs (Ashbrook et al., 2022; Li et al., 2020; McLaren et al., 2016).

Candidate analysis identifies *Smarcal1* and *Usp37* as potential effectors of PGB burden

To find the allele variants most likely to affect PGB density from within the *Pgb1* locus, we devised specific selection criteria (**Figure 3A**). We only considered protein coding genes with a linkage –logP ≥ 3 (n = 36 protein coding genes), as this increased the likelihood of identifying causal gene variants. It is worth noting that allele variants in gene candidates could influence PGB density via two distinct mechanisms, either by altering expression levels or by modifying protein function. In order to identify genes that exhibit altered expression levels due to allele variations within our target locus (a *cis*-regulatory effect), we collected hippocampal exon array data from 235 BXD animals (223–972 days old, GN dataset ID: GN392). We conducted mRNA expression-QTL analysis to identify the locus controlling expression level differences among BXD strains. Notably, we found that seven out of 36 candidate genes mapped as a *cis*-eQTL to the *Pgb1* locus (–logP ≥ 1.5). These *cis*-eQTL peaks suggested that potential gene variants may be affecting promoter or enhancer regions in close proximity to the candidate gene loci. We selected those candidates that displayed a significant correlation (p < 0.05) between transcript expression levels and PGB burden and identified three candidate, namely *Smarcal1, Tns1,* and *Usp37* (**Figure 3B**). *Smarcal1 and Tns1 were* negatively correlated with PGB burden—correlations of –0.73 and –0.52, respectively—indicating a potential protective role for these genes against PGB burden. In contrast, *Usp37* exhibited a positive correlation with PGB burden (r = 0.509), suggesting a potential role in enhancing PGB numbers. Taken together, our analysis led to the identification of three candidate genes that may impact PGB burden through changes in gene expression levels.

To identify candidates that may affect PGB burden via changes in protein function, we focused on candidates with at least one non-synonymous mutation between the *B* and *D* alleles. Out of the eight candidate genes that carried such mutations, we selected four candidates carrying variants with a previously reported high probability of affecting protein function (Wang et al., 2016): *Smarcal1*, *Usp37*, *Zfp142*, and *Stk36* (**Figure 3C**). We generated 3D protein structures of these candidates to visualize the precise location of potentially high-impact gene variants (**Figure 3D**). Interestingly, two candidates may affect PGB density through both changes in expression levels as well as through changes in protein function; the annealing helicase *Smarcal1* and the deubiquitinase *Usp37*.

Phenome-wide association reveals *trans*-regulation of *Hp1bp3* mRNA expression by *Pgb1*

We next evaluated whether the *Pgb1* locus has a *trans*-acting regulatory effect on the expression levels of mRNA or proteins from genes located in different regions of the genome. To this end, we performed a phenome wide association study (PheWAS) using SNP rs31200269 as a surrogate marker for the *Pgb1* locus, since this SNP marker has the highest association with PGB density. This SNP was located on Chr 1 at 72.626982 Mb (mm10 assembly) in intron 14 of *Smarcal1*.

First, we applied a PheWAS test with the aged hippocampal exon array data generated for 67 BXD strains and 234 mice which matched the age at which PGB burden was measured (mean age of 18 months). The strongest *trans*-acting regulatory effect of *Pgb1* was found with the expression of the terminal exon of heterochromatin protein 1 binding protein 3 (*Hp1bp3*) (probe set 10340721) (**Figure 4A,B** and **Table 1A**). *Hp1bp3* is located on Chr 4, and is a modulator of cognitive function and conditioned fear memory in older BXD mice (Neuner et al., 2016). *Hp1bp3* expression in the aged hippocampus mapped to the *Pgb1* locus with a –logP of 6.3, showing higher expression in the *D* allele compared to the *B* allele (**Figure 4C**). PGB burden and *Hp1bp3* expression covaried negatively (r = –0.38; p = 0.055, Spearman correlation) with 26 strains measured in common. Expression of *Hp1bp3* increased as a function of age in a way consistent with linkage to PGB burden (**Figure 4D**, r = 0.29, $p < 1 \times 10^{-5}$) (Jucker et al., 1992).

Next, we performed a matched PheWAS of an aged hippocampal proteome data generated for 205 BXD females (6–24 months of age) (sample details as in (Williams et al., 2022)). We surveyed 17,799 proteins to identify those being *trans*-regulated by the *Pgb1* locus. A total of six proteins, including GSTP2, UBA1, and NPEPPS showed linkage scores at *Pgb1* with –logP ≥ 2 (**Table 1B**). Taken together, we identified nine transcripts and six proteins to be *trans*-regulated by the *Pgb1* locus.

Phenome-wide association and conditioned fear memory testing suggest a limited effect of PGB burden on cognitive function

To assess the impact of PGB burden on cognition, we investigated phenome-wide associations between *Pgb1* and any of approximately 12,000 phenotypes available at genenetwork.org. Notably, only two behavioral traits were linked to the *Pgb1* locus, including the level of activity in a Y-maze in 8-month-old females (r = –0.78, GN trait ID: BXD_20820) (Neuner et al., 2019). This trait mapped to Chr 1 with a score of –logP of 4.8 at the *Pgb1* locus, but considering that the peak lay outside the *Pgb1* locus, the trait likely mapped due to an

upstream allele variant and would therefore not be associated with PGB burden. The *Pgb1* marker also covaried with cocaine-induced stereotypy (chewing, rearing, GN trait ID: BXD_10315) in young adults (r = 0.57, n = 23) with a linkage score of 4.2 and a high *D* allele (Jones et al., 1999), which is not directly linked with cognition. Hence, we did not observe convincing evidence of an effector role for *Pgb1* or PGBs on cognitive performance.

Even though we analyzed ~12,000 phenotypes, only limited behavioral data is available for aged animals. To mitigate this limitation, we performed conditioned fear memory (CFM) tests on aged BXDs. We produced a dataset of conditioned fear memory (GN trait ID: 5XF_10474) with 14 month old B-BXD animals from 42 different strains (**Figure 5A**). At this age, PGBs are present in the B6 hippocampus and age-related behavioral alterations such as reduced CFM become apparent (Amelchenko et al., 2023; Hamieh et al., 2021; Jucker et al., 1992; Wu et al., 2015). Importantly, no correlation was found between CFM (time spent frozen) and PGB density in the subset of overlapping strains (n = 13) (**Figure 5B,C**). Interestingly, a significant correlation ($p = 5.18 \times 10^{-3}$) was found between *Pgb1* and CFM when considering all 134 animals across 42 strains, showing that animals with a *Pgb1 BB* allele outperforms those with a *Pgb1 BD* allele (**Figure 5D**). Considering that animals carrying a *Pgb1 BB* allele both have higher PGB burden as well as better CFM performance, these data suggest that it is unlikely that PGB burden has a negative impact on cognition. Therefore, we conclude it is unlikely that age-related PGB burden in the hippocampus has a strong negative effect on cognitive function.

Discussion

The accumulation of polyglucosan bodies in astrocytes is linked to aging and disease (Duran, 2023; Duran et al., 2014). However, the mechanisms by which PGBs are formed, as well as their clinical relevance, remain poorly understood. Here, we harnessed natural variation of hippocampal PGB density that segregates in the BXD family to map the underlying genetic causes of these aggregates and to nominate candidate genes that potentially influence this age-related phenotype. Our work provides a foundation for investigation into identifying molecular mechanisms modulating this process. In addition, we investigated the potential relevance of PGBs on age-related cognitive decline.

PGBs in the aging mouse brain have first been described more than 40 years ago (Akiyama et al., 1986; Lamar et al., 1976; Mandybur et al., 1989). However, they did not receive significant attention until they were mistakenly assumed to be amyloid-lesions in mouse models of Alzheimer's disease (Baglietto-Vargas et al., 2021; Jucker et al., 1992). Subsequent studies (Jucker et al., 1994a) revealed that PGB abundance and onset of development have a substantial heritable component, with abundant lesions and early onset in some strains of mice (e.g., C57BL/6, SAMP8, AKR/J) and rare lesions and late onset in other strains of mice (e.g., A, DBA/2J, CBA/J, C3H/HeJ). We validated that polyglucosan density in the hippocampus varies widely among BXD strains. Using genenetwork.org and QTL mapping, we provided evidence that variation of polyglucosan body occurrence in the BXD family mapped to a region on chromosome 1 (72–75 Mb) with a highly significant correlation (–logP = 4.9). In total, 47 coding transcripts were located within this *Pgb1* locus, from which we identified two protein coding candidates likely to impact PGB burden, *Smarcal1* and *Usp37*. Both candidates were known to harbor a non-synonymous mutation with the potential to impact protein function, as well as a significant correlation between their mRNA expression

levels and PGB burden. *Smarcal1* encodes an annealing helicase involved in repairing forked ends of DNA in response to DNA damage (Bansal et al., 2020) and *Usp37* encodes a deubiquitinase with limited functional characterisation, although interestingly may also be linked to the DNA damage response (Wu et al., 2021). *Smarcal1* and *Usp37* have not been directly linked to glycogen metabolism. Hence, follow-up studies should focus on testing the potential mechanistic link between these candidate genes and PGB formation and degradation, especially due to the involvement of *Usp37* in the ubiquitin system and the link between both *Smarcal1* and *Usp37* with DNA damage, a key hallmark of aging (Schumacher et al., 2021).

Phenome-wide association to the *Pgb1* locus linked *Hp1bp3* with PGB density. Interestingly, *Hp1bp3* has been shown to be a key regulator of cognitive function and conditioned fear memory in older BXD mice (Neuner et al., 2016). Hence, the *Pgb1*-mediated upregulation of *Hp1bp3* might prevent animals with a *D* allele from developing a high PGB density. Despite the potential link between *Hp1bp3* and the *Pgb1* locus and PGB density, no convincing evidence was found that directly links any behavioral trait with the *Pgb1* locus or PGB burden. Although around 12,000 phenotypes were analyzed for a potential correlation with *Pgb1*, only limited behavioral data is available for aged BXD mice. To mitigate this limitation, we performed conditioned fear memory (CFM) tests on aged B-BXDs, which again yielded no evidence suggesting negative effects of PGBs on cognition. Studies have shown that CFM is a effective readout of hippocampal function in spatial memory and learning confidence (Amelchenko et al., 2023; Hernández-Mercado and Zepeda, 2022) and have been demonstrated to deliver readable outcomes in cognitive performance tests in an age related context (Hamieh et al., 2021; Hernández-Mercado and Zepeda, 2022; Verbitsky et al., 2004; Wu et al., 2015). Although we cannot exclude an effect of PGBs on cognition at geriatric stages, we conclude it is unlikely that age-related PGB burden in the hippocampus has a strong negative effect on cognitive function up to 18-months of age.

Recent work by *Lee et al.* (2022) has identified an age-related accumulation of autophagy-dysregulated astrocytes (APDAs) in the hippocampus of mice, which negatively affect synapse number and homeostasis of adjacent neurons (Lee et al., 2022). Future research should investigate whether the astrocytes containing the PGBs we identify here are in fact the same astrocytes as those termed APDAs, or whether these are two distinct astrocyte subtypes.

Taken together, we have identified a novel modifier locus controlling PGB density in the hippocampus and shed light on the genetic architecture and clinical relevance of this strikingly diverse hippocampal phenotype.

Materials and methods

Animals and tissue preparation protocol for PGB assessment

Mice used for the assessment of PGBs were obtained as retired breeders between 1997 and 1999 from The Jackson Laboratory (Bar Harbor) and further aged at the Gerontology Research Center, NIA, NIH. A total of 144 female mice (all 17-19 months of age) were used: 32 BXD strains (n=133; 2 to 10 mice/strain; mean 4.1/strain), C57BL/6J (n=4) and DBA/2J (n=5). See supplementary Table 1 for details. BXDs are progeny of crosses of female C57BL/6J (B6 or B) and male DBA/2J (D2 or D) parents. The BXD family was founded in the

early 1970s by Taylor (Taylor et al., 1999) and was expanded in multiple epochs from ~80 to over 140 strains (Ashbrook et al., 2021; Peirce et al., 2004). Animals were euthanized by pentobarbital overdose and perfused transcardially with 0.1 M phosphate-buffered saline (PBS) at pH 7.4. Brains were removed and immersion fixed in 4% paraformaldehyde, followed by 30% sucrose, and freezing in isopentane at -20°C. Coronal sections throughout the hippocampus were cut on a freezing-sliding microtome at 25 µm. Visualisation of the polyglucosan bodies (PGB) was done as previously described (Jucker et al., 1994b). In short, free-floating sections were rinsed with TBS (0.05 M Tris-buffer containing 1.5% NaC1, pH 7.4) and subsequently incubated in 0.3% Triton X-100 in TBS, followed by an incubation in 5% of goat serum in TBS. Sections were reacted for two days at 4°C with an antibody against a laminin-binding protein (LBP110) in TBS containing 2% blocking serum and 0.3% Triton X-100. Secondary antibody was goat anti-rabbit IgG. Antibodies were then detected by the ABC method with reagents from Vector Laboratories (Vectastain Elite ABC Kit, Vector Laboratories, Burlingame, CA, U.S.A.). The chromogen was diaminobenzidine, and the reaction product was intensified by adding NiCl.

Quantification of PGBs

PGB are typically detected with periodic acid-Schiff staining (Jucker et al., 1994a, 1992; Manich et al., 2016). However, an antibody raised against the laminin-binding protein 110 (LBP110) was found to strongly bind PGB-lesions at light- and ultrastructural level and is largely used to identify PBGs in histological preparations (Jucker et al., 1994a, 1992; Kuo et al., 1996). The number of PGBs was assessed in a random systematic sampled set of every 10th section through the entire hippocampus. Total number of clusters of LBP110-positive PGB per unilateral hippocampus was estimated by counting all clusters PGB in all sections and multiplying the number with the section sample fraction (10 for the present analysis). The PGB numbers among strains was strongly skewed and values were therefore $\log_{10}$ transformed prior to mapping (see Table S1).

Hippocampal transcriptome in older BXD animals (GN392)

Animals. 234 individual mice (140 females and 94 males) from 71 BXD strains, parental C57BL/6J, DBA/2J and D2B6F1 hybrid, were used in producing the hippocampal transcriptome dataset (GN392). These mice were obtained from UTHSC, ORNL, Beth Israel Deaconess or directly from The Jackson. The age for a great majority of mice was between 12 and 28 months (average of 18 months). The animals were sacrificed under saturated isoflurane and brains were removed and placed in RNAlater prior to dissection. Cerebella and olfactory bulbs were removed; brains were hemisected, and both hippocampi were dissected whole in Dr. Lu's lab. Hippocampal samples are close to complete (see Lu et al., 2001) yet may include variable amounts of subiculum and fimbria. All procedures were approved by the UTHSC Institutional Animal Care and Use Committee.

RNA Extraction and data generation. RNA was extracted using the RNeasy mini kit (Qiagen, Valencia, CA, USA) according to the manufactures' procedure. RNA purity and concentration was checked using 260/280 nm absorbance ratio and RNA integrity was analyzed using the Agilent Bioanalyzer 2100 (Agilent Technologies). Samples with RNA Integrity Numbers (RIN values) > 7.5 were run on Affy MoGene1.0 ST at the UTHSC. This array contains 34,700 probe sets that target ~29,000 well-defined transcripts (RefSeq mRNA isoforms). A single probe set is a collection of about 27 probes that target known exons within a single gene. The

multiple probes design provides a more comprehensive coverage of transcripts from a single gene.

Data processing. Probe set level intensity values were extracted from the CEL files using the Affymetrix GeneChip Operating Software. Data normalization was performed using the R package "Affy" available from www.Bioconductor.org. The Robust Multichip Averaging protocol was used to process the expression values. The array data was then log transformed and rescaled using a z-scoring procedure to set the mean of each sample at eight expression units with a SD of 2 units.

Quantitative trait mapping

Two datasets were generated. The original untransformed dataset for 32 members of the BXD family (GN trait BXD_10685) has a range of values from almost free of any polyglucosan aggregates (1.5 PGBs per hippocampus in BXD32) to very high densities of PGB (646 PGBs per hippocampus in BXD20). Due to the extreme skew (3.3) and kurtosis (12) the original data are not well suited for mapping (Figure S1). In contrast, $\log_{10}$ transformed data (GN trait BXD_10686) were suitable for mapping and have low skew (0.29) and kurtosis (–0.58), and mean and median are both about 1.4. (Figure 1D). We used new whole-genome sequence (WGS) marker maps (Ashbrook et al., 2021; Sasani et al., 2022) and the GRCm38 (mm10) mouse assembly. These high density genetic maps were used in combination with GEMMA, a linear mixed model method which corrects for differences in relatedness among strains (Zhou and Stephens, 2012). We mapped using a set of 20 kinship matrixes, each computed by leaving out all markers on one of the 19 autosomes or the X chromosome (Yang et al., 2014). Support intervals are given by the –2.0 logP drop from the high point. Effect sizes are uncorrected for the Beavis effect but do not include parental phenotypes (Beavis, 1994; Beavis et al., 1991; Xu, 2003).

Selection of data on candidate genes shown in Table S1

In Table S1, we list key data on coding genes within the *Pgb1* locus.

Mapping mRNA expression. We mapped every gene on the locus by itself from the "UTHSC BXD Aged Hippocampus Affy Mouse Gene 1.0 ST(Sep 12) RMA Exon Level" dataset provided in genenetwork.org (GN Accession: GN392), choosing between multiple results per gene based based on the amount of represented strains overlapping with the PGB strains, standard error being low, effect sizes being high and highest possible association scores. We noted the mean mRNA expression, given as a normalized Z-like scores value with a mean value of 8 with a SD of two (Geisert et al., 2009) in hippocampal array data, to confirm the gene being represented in our tissue of interest. Values lower than 7 were considered to represent negligible expression. *cis*-eQTLs using microarray data can contain false positives due to higher binding affinity of the B allele, for which the microarray probes were designed. We reduced the possibility of the *cis-eQTL* being a probe binding artifact, by verifying its existence in other tissues as well, following the assumption of the *cis-eQTL* effect being predominantly conserved over tissues (Westra and Franke, 2014). Furthermore, we have prioritized candidates without proven irregularities in array sets found in mismatched SNP alleles on the UCSC Genome Browser on Mouse (GRCm38/mm10), another redirection given by genenetwork.org to verify probes via the "BLAT" tool (Kent, 2002).

Identification of allele variations affecting protein function. Proteomics data on the BXD family provided by David Ashbrook (Ashbrook et al., 2022) was used to investigate predicted protein altering segregating variants and whether those would be predicted to be deleterious variants

according to (McLaren et al., 2016), providing another category measuring the severity of consequences of a gene variation in the family. Subsequently, we examined genes with a high number of mutations based on (Wang et al., 2016), which indicates the mutation severity based on the maximum Grantham score.

Age associated gene expression changes. Age-associated gene expression changes were quantified using bulk RNA-seq data from C57BL/6J male mouse hippocampi across the lifespan (2 months, 12 months, 24 months, n=3 per timepoint; (Li et al., 2020)). Raw paired end sequencing reads were quality assessed and processed through fastp (v0.20.1; Phred quality > 40; 10% unqualified bases permitted) (Chen et al., 2018). Processed reads were mapped and quantified using Salmon (v1.10.1) in mapping-based mode with seqBias, gcBias and posBias parameters enabled (Patro et al., 2017). A decoy-aware transcriptome assembly (concatenated genome and transcriptome) used for mapping was derived using Mus musculus GRCm38 reference files (release 102 from Ensembl). Gene level expression estimates were achieved using tximport (v1.22.0) (Soneson et al., 2015) and normalized counts calculated using DESeq2 (v.1.34.0) (Love et al., 2014).

Assessing cell type specific expression. Single-cell RNA-seq data was retrieved from the Allen Brain Atlas portal, to assess cell-type specific expression of the candidate genes (Yao et al., 2021).

PGB burden impact in cognitive performance

To account for the potential influence of accumulation of PGBs in hippocampal astrocytes on learning and memory functions, a search was proposed using the dataset: “UTHSC BXD Aged Hippocampus Affy Mouse Gene 1.0 ST (Sep12) RMA Exon Level” (GN Accession: GN392). The search was "rif=learning", where rif = reference into function—an NCBI summary of gene functions with links to references was used. The list of results was sorted by peak location and focused on the genes that map on 72–75 Mb. Finally, mapping was repeated using GEMMA to validate the results.

Phenome-wide association analysis of *Pgb1* locus

Two SNP markers within *Pgb1* were selected—a proximal marker, rs31200269, at 72.62 Mb and a distal marker, rs13475923, at 74.07 Mb. These markers are in strong linkage (r^2 = 0.79) and only two strains among those phenotyped differ in their genotypes—BXD23 and BXD29. Both proximal and distal markers were used as proxies for the *Pgb1* locus and in the phenome-wide association analysis to find traits that a) covary with these markers and b) that map within the *Pgb1* locus with linkage scores (–log*P* values) higher than 2.0. Candidate phenotypes, including expression traits, were considered of high interest when they were associated with behavioral differences or with the metabolism of polyglucosan or the formation of other types of aggregates in CNS.

Contextual Fear Conditioning

Standard contextual fear conditioning was collected as has been used by our group and others previously (Neuner et al., 2019). This was used to characterize cognitive function across the B-BXDs at either 6 or 14 months of age. We used F1 crosses between the C57BL/6J and BXD strains (B6-BXDF1s), as these were being phenotyped in a parallel project. Krass et al., (2003) showed that C57BL/6J animals have a high lesion density (mean 19 lesions/mm2), while DBA/2J and B6xD2F1s have very low lesion density (mean 0.04 and 0.65 lesions/mm2). This

indicates that the effect of the causal C57BL/6J genotype is recessive, requiring two copies of the B6-allele. In the current study using F1 crosses between the C57BL/6J and BXD strains, approximately half the population will have the BB (high risk) genotype at any position, and half will have a BD (low risk) genotype. On the first day of training, mice were placed in a training chamber and four 0.9 mA 1 s foot shocks were delivered after a baseline period of 180 s and then repeatedly after an interchangeable interval of 115 ± 20 s. Four post- shock intervals were defined as the 40 s following the offset of each foot (Colbourn Instruments, PA, United States). The percentage of time spent freezing following the final shock was used as a measure of contextual fear acquisition across the panel. Twenty- four hours after training, mice were placed back into the training chamber and the percentage of time spent freezing throughout the entire 10-min test was measured as an index of contextual fear memory. A DeepLabCut model was trained to recognize 13 points on a mouse, following the labelling system of (Sturman et al., 2020). DLCAnalyzer was used to measure freezing events (Sturman et al., 2020).

Data Availability

All used data is publicly available. The overview of the used datasets and where to find them can be found in *online supplementary table 2 (Table S2).*

Acknowledgements

We would like to thank the Summer School *Systems Genetics of Neural Ageing* for bringing us together and spurring our international collaboration. We would also like to acknowledge the funding for the Summer School 2022 from the e:Med Systems Medicine programme of the BMBF (*Bundesministerium für Bildung und Forschung*; German Ministry of Education and Research) to RWO. Alicia Gómez-Pascual is supported by Fundación Séneca, Región de Murcia, Spain (21259/FPI/19). D.E.M. de Bakker is financed by a Rubicon scholarship (452021116) from the Dutch Research Council (NWO). NIH NIA R01AG070913-01 (Williams/Johnson), R01AG075813-01 (Ashbrook) and R01AG075818 (Kaczorowski). This research was funded in whole or in part by Aligning Science Across Parkinson's [ASAP-000509] through the Michael J. Fox Foundation for Parkinson's Research (MJFF), which supported co-author J.E.T (ORCID: 0000-0003-3010-6634).

References

Akiyama H, Kameyama M, Akiguchi I, Sugiyama H, Kawamata T, Fukuyama H, Kimura H, Matsushita M, Takeda T. 1986. Periodic acid-Schiff (PAS)-positive, granular structures increase in the brain of senescence accelerated mouse (SAM). *Acta Neuropathol* **72**:124–129. doi:10.1007/BF00685973

Amelchenko EM, Bezriadnov DV, Chekhov OA, Anokhin KV, Lazutkin AA, Enikolopov G. 2023. Age-related decline in cognitive flexibility is associated with the levels of hippocampal neurogenesis. *Frontiers in Neuroscience* **17**. doi:10.3389/fnins.2023.1232670

Ashbrook DG, Arends D, Prins P, Mulligan MK, Roy S, Williams EG, Lutz CM, Valenzuela A, Bohl CJ, Ingels JF, McCarty MS, Centeno AG, Hager R, Auwerx J, Lu L, Williams RW. 2021. A platform for experimental precision medicine: The extended BXD mouse family. *Cell Syst* **12**:235-247.e9. doi:10.1016/j.cels.2020.12.002

Ashbrook DG, Sasani T, Maksimov M, Gunturkun MH, Ma N, Villani F, Ren Y, Rothschild D, Chen H, Lu L, Colonna V, Dumont B, Harris K, Gymrek M, Pritchard JK, Palmer AA, Williams RW. 2022. Private and sub-family specific mutations of founder haplotypes in the BXD family reveal phenotypic consequences relevant to health and disease. doi:10.1101/2022.04.21.489063

Baglietto-Vargas D, Forner S, Cai L, Martini AC, Trujillo-Estrada L, Swarup V, Nguyen MMT, Do Huynh K, Javonillo DI, Tran KM, Phan J, Jiang S, Kramár EA, Nuñez-Diaz C, Balderrama-Gutierrez G, Garcia F, Childs J, Rodriguez-Ortiz CJ, Garcia-Leon JA, Kitazawa M, Shahnawaz M, Matheos DP, Ma X, Da Cunha C, Walls KC, Ager RR, Soto C, Gutierrez A, Moreno-Gonzalez I, Mortazavi A, Tenner AJ, MacGregor GR, Wood M, Green KN, LaFerla FM. 2021. Generation of a humanized Aβ expressing mouse demonstrating aspects of Alzheimer's disease-like pathology. *Nat Commun* **12**:2421. doi:10.1038/s41467-021-22624-z

Bansal R, Hussain S, Chanana UB, Bisht D, Goel I, Muthuswami R. 2020. SMARCAL1, the annealing helicase and the transcriptional co-regulator. *IUBMB Life* **72**:2080–2096. doi:10.1002/iub.2354

Beavis WD. 1994. The power and deceit of QTL experiments: lessons from comparative QTL studiesProceedings of the Forty-Ninth Annual Corn and Sorghum Industry Research Conference. Chicago, IL. p. 266.

Beavis WD, Grant D, Albertsen M, Fincher R. 1991. Quantitative trait loci for plant height in four maize populations and their associations with qualitative genetic loci. *Theor Appl Genet* **83**:141–145. doi:10.1007/BF00226242

Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**:i884–i890. doi:10.1093/bioinformatics/bty560

Doron A, Rubin A, Benmelech-Chovav A, Benaim N, Carmi T, Refaeli R, Novick N, Kreisel T, Ziv Y, Goshen I. 2022. Hippocampal astrocytes encode reward location. *Nature* **609**:772–778. doi:10.1038/s41586-022-05146-6

Duran J. 2023. Role of Astrocytes in the Pathophysiology of Lafora Disease and Other Glycogen Storage Disorders. *Cells* **12**:722. doi:10.3390/cells12050722

Duran J, Gruart A, Garcia-Rocha M, Delgado-Garcia JM, Guinovart JJ. 2014. Glycogen accumulation underlies neurodegeneration and autophagy impairment in Lafora disease. *Human Molecular Genetics* **23**:3147–3156. doi:10.1093/hmg/ddu024

Duran J, Guinovart JJ. 2015. Brain glycogen in health and disease. *Molecular Aspects of Medicine* **46**:70–77. doi:10.1016/j.mam.2015.08.007

Gautier L, Cope L, Bolstad BM, Irizarry RA. 2004. affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**:307–315. doi:10.1093/bioinformatics/btg405

Geisert EE, Lu L, Freeman-Anderson NE, Templeton JP, Nassr M, Wang X, Gu W, Jiao Y, Williams RW. 2009. Gene expression in the mouse eye: an online resource for genetics using 103 strains of mice. *Molecular Vision*.

Hamieh AM, Camperos E, Hernier AM, Castagné V. 2021. C57BL/6 mice as a preclinical model to study age-related cognitive deficits: Executive functions impairment and inter-individual differences. *Brain Research* **1751**:147173. doi:10.1016/j.brainres.2020.147173

Hernández-Mercado K, Zepeda A. 2022. Morris Water Maze and Contextual Fear Conditioning Tasks to Evaluate Cognitive Functions Associated With Adult Hippocampal Neurogenesis. *Frontiers in Neuroscience* **15**. doi:10.3389/fnins.2021.782947

Jones BC, Tarantino LM, Rodriguez LA, Reed CL, McClearn GE, Plomin R, Erwin VG. 1999. Quantitative-trait loci analysis of cocaine-related behaviours and neurochemistry. *Pharmacogenetics* **9**:607–617.

Jucker M, Walker LC, Kuo H, Tian M, Ingram DK. 1994a. Age-related fibrillar deposits in brains of C57BL/6 mice. A review of localization, staining characteristics, and strain specificity. *Mol Neurobiol* **9**:125–133. doi:10.1007/BF02816112

Jucker M, Walker LC, Martin LJ, Kitt CA, Kleinman HK, Ingram DK, Price DL. 1992. Age-associated inclusions in normal and transgenic mouse brain. *Science* **255**:1443–1445. doi:10.1126/science.1542796

Jucker M, Walker LC, Schwarb P, Hengemihle J, Kuo H, Snow AD, Bamert F, Ingram DK. 1994b. Age-related deposition of glia-associated fibrillar material in brains of C57BL/6 mice. *Neuroscience* **60**:875–889. doi:10.1016/0306-4522(94)90269-0

Kent WJ. 2002. BLAT—The BLAST-Like Alignment Tool. *Genome Res* **12**:656–664. doi:10.1101/gr.229202

Krass, K. L., Colinayo, V., Ghazalpour, A., Vinters, H. V., Lusis, A. J., & Drake, T. A. (2003). Genetic loci contributing to age-related hippocampal lesions in mice. Neurobiology of disease, 13(2), 102–108. https://doi.org/10.1016/s0969-9961(03)00034-2

Kuo H, Ingram DK, Walker LC, Tian M, Hengemihle JM, Jucker M. 1996. Similarities in the age-related hippocampal deposition of periodic acid-schiff-positive granules in the senescence-accelerated mouse P8 and C57BL/6 mouse strains. *Neuroscience* **74**:733–740. doi:10.1016/0306-4522(96)00169-8

Lamar CH, Hinsman EJ, Henrikson CK. 1976. Alterations in the hippocampus of aged mice. *Acta Neuropathol* **36**:387–391. doi:10.1007/BF00699644

Lee E, Jung Y-J, Park YR, Lim S, Choi Y-J, Lee SY, Kim CH, Mun JY, Chung W-S. 2022. A distinct astrocyte subtype in the aging mouse brain characterized by impaired protein homeostasis. *Nat Aging* **2**:726–741. doi:10.1038/s43587-022-00257-1

Li M, Su S, Cai W, Cao J, Miao X, Zang W, Gao S, Xu Y, Yang J, Tao Y-X, Ai Y. 2020. Differentially Expressed Genes in the Brain of Aging Mice With Cognitive Alteration and Depression- and Anxiety-Like Behaviors. *Frontiers in Cell and Developmental Biology* **8**.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**:550. doi:10.1186/s13059-014-0550-8

Lu L, Airey DC, Williams RW. 2001. Complex Trait Analysis of the Hippocampus: Mapping and Biometric Analysis of Two Novel Gene Loci with Specific Effects on Hippocampal Structure in Mice. *J Neurosci* **21**:3503–3514. doi:10.1523/JNEUROSCI.21-10-03503.2001

Mandybur TI, Ormsby I, Zemlan FP. 1989. Cerebral aging: a quantitative study of gliosis in old nude mice. *Acta Neuropathol* **77**:507–513. doi:10.1007/BF00687252

Manich G, Cabezón I, Augé E, Pelegrí C, Vilaplana J. 2016. Periodic acid-Schiff granules in the brain of aged mice: From amyloid aggregates to degenerative structures containing neo-epitopes. *Ageing Res Rev* **27**:42–55. doi:10.1016/j.arr.2016.03.001

McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, Flicek P, Cunningham F. 2016. The Ensembl Variant Effect Predictor. *Genome Biology* **17**:122. doi:10.1186/s13059-016-0974-4

Moreno-Estellés M, Campos-Rodríguez Á, Rubio-Villena C, Kumarasinghe L, Garcia-Gimeno MA, Sanz P. 2023. Deciphering the Polyglucosan Accumulation Present in Lafora Disease Using an Astrocytic Cellular Model. *Int J Mol Sci* **24**:6020. doi:10.3390/ijms24076020

Neuner SM, Garfinkel BP, Wilmott LA, Ignatowska-Jankowska BM, Citri A, Orly J, Lu L, Overall RW, Mulligan MK, Kempermann G, Williams RW, O'Connell KMS, Kaczorowski CC. 2016. Systems genetics identifies Hp1bp3 as a novel modulator of cognitive aging. *Neurobiol Aging* **46**:58–67. doi:10.1016/j.neurobiolaging.2016.06.008

Neuner SM, Heuer SE, Huentelman MJ, O'Connell KMS, Kaczorowski CC. 2019. Harnessing Genetic Complexity to Enhance Translatability of Alzheimer's Disease Mouse Models: A Path toward Precision Medicine. *Neuron* **101**:399-411.e5. doi:10.1016/j.neuron.2018.11.040

Nitschke S, Sullivan MA, Mitra S, Marchioni CR, Lee JPY, Smith BH, Ahonen S, Wu J, Chown EE, Wang P, Petković S, Zhao X, DiGiovanni LF, Perri AM, Israelian L, Grossman TR, Kordasiewicz H, Vilaplana F, Iwai K, Nitschke F, Minassian BA. 2022. Glycogen synthase downregulation rescues the amylopectinosis of murine RBCK1 deficiency. *Brain* **145**:2361–2377. doi:10.1093/brain/awac017

Palmer AL, Ousman SS. 2018. Astrocytes and Aging. *Front Aging Neurosci* **10**:337. doi:10.3389/fnagi.2018.00337

Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. 2017. Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods* **14**:417–419. doi:10.1038/nmeth.4197

Peirce JL, Lu L, Gu J, Silver LM, Williams RW. 2004. A new set of BXD recombinant inbred lines from advanced intercross populations in mice. *BMC Genet* **5**:7. doi:10.1186/1471-2156-5-7

Sasani TA, Ashbrook DG, Beichman AC, Lu L, Palmer AA, Williams RW, Pritchard JK, Harris K. 2022. A natural mutator allele shapes mutation spectrum variation in mice. *Nature* **605**:497–502. doi:10.1038/s41586-022-04701-5

Schumacher B, Pothof J, Vijg J, Hoeijmakers JHJ. 2021. The central role of DNA damage in the ageing process. *Nature* **592**:695–703. doi:10.1038/s41586-021-03307-7

Soneson C, Love M, Robinson M. 2015. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences [version 1; peer review: 2 approved]. *F1000Research* **4**. doi:10.12688/f1000research.7563.1

Sturman O, von Ziegler L, Schläppi C, Akyol F, Privitera M, Slominski D, Grimm C, Thieren L, Zerbi V, Grewe B, Bohacek J. 2020. Deep learning-based behavioral analysis reaches human accuracy and is capable of outperforming commercial solutions. *Neuropsychopharmacol* **45**:1942–1952. doi:10.1038/s41386-020-0776-y

Taylor BA. 1978. Recombinant inbred strains: use in gene mapping. *Origins of Inbred Mice* **New York: Academic Press**:423–438. doi:10.1016/B978-0-12-507850-4.50032-9

Taylor BA, Wnek C, Kotlus BS, Roemer N, MacTaggart T, Phillips SJ. 1999. Genotyping new BXD recombinant inbred mouse strains and comparison of BXD and consensus maps. *Mammalian Genome* **10**:335–348. doi:10.1007/s003359900998

Verbitsky M, Yonan AL, Malleret G, Kandel ER, Gilliam TC, Pavlidis P. 2004. Altered Hippocampal Transcript Profile Accompanies an Age-Related Spatial Memory Deficit in Mice. *Learn Mem* **11**:253–260. doi:10.1101/lm.68204

Viana JF, Machado JL, Abreu DS, Veiga A, Barsanti S, Tavares G, Martins M, Sardinha VM, Guerra-Gomes S, Domingos C, Pauletti A, Wahis J, Liu C, Calì C, Henneberger C, Holt MG, Oliveira JF. 2023. Astrocyte structural heterogeneity in the mouse hippocampus. *Glia* **71**:1667–1682. doi:10.1002/glia.24362

Wang W, Lohi H, Skurat AV, DePaoli-Roach AA, Minassian BA, Roach PJ. 2007. Glycogen metabolism in tissues from a mouse model of Lafora disease. *Archives of biochemistry and biophysics* **457**:264. doi:10.1016/j.abb.2006.10.017

Wang X, Pandey AK, Mulligan MK, Williams EG, Mozhui K, Li Z, Jovaisaite V, Quarles LD, Xiao Z, Huang J, Capra JA, Chen Z, Taylor WL, Bastarache L, Niu X, Pollard KS, Ciobanu DC, Reznik AO, Tishkov AV, Zhulin IB, Peng J, Nelson SF, Denny JC, Auwerx J, Lu L, Williams RW. 2016. Joint mouse–human phenome-wide association to test gene function and disease risk. *Nat Commun* **7**:10464. doi:10.1038/ncomms10464

Westra H-J, Franke L. 2014. From genome to function by studying eQTLs. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, From genome to function **1842**:1896–1902. doi:10.1016/j.bbadis.2014.04.024

Williams EG, Pfister N, Roy S, Statzer C, Haverty J, Ingels J, Bohl C, Hasan M, Čuklina J, Bühlmann P, Zamboni N, Lu L, Ewald CY, Williams RW, Aebersold R. 2022. Multiomic profiling of the liver across diets and age in a diverse mouse population. *Cell Syst* **13**:43-57.e6. doi:10.1016/j.cels.2021.09.005

Wu C, Chang Y, Chen J, Su Y, Li L, Chen Y, Li Y, Wu J, Huang J, Zhao F, Wang W, Yin H, Wang S, Jin M, Lou Z, Zhu W-G, Luo K, Zhang J, Yuan J. 2021. USP37 regulates DNA damage response through stabilizing and deubiquitinating BLM. *Nucleic Acids Res* **49**:11224–11240. doi:10.1093/nar/gkab842

Wu MV, Luna VM, Hen R. 2015. Running rescues a fear-based contextual discrimination deficit in aged mice. *Front Syst Neurosci* **9**. doi:10.3389/fnsys.2015.00114

Xu S. 2003. Theoretical basis of the Beavis effect. *Genetics* **165**:2259–2268. doi:10.1093/genetics/165.4.2259

Yang J, Zaitlen NA, Goddard ME, Visscher PM, Price AL. 2014. Advantages and pitfalls in the application of mixed-model association methods. *Nat Genet* **46**:100–106. doi:10.1038/ng.2876

Yao Z, van Velthoven CTJ, Nguyen TN, Goldy J, Sedeno-Cortes AE, Baftizadeh F, Bertagnolli D, Casper T, Chiang M, Crichton K, Ding S-L, Fong O, Garren E, Glandon A, Gouwens NW, Gray J, Graybuck LT, Hawrylycz MJ, Hirschstein D, Kroll M, Lathia K, Lee C, Levi B, McMillen D, Mok S, Pham T, Ren Q, Rimorin C, Shapovalova N, Sulc J, Sunkin SM, Tieu M, Torkelson A, Tung H, Ward K, Dee N, Smith KA, Tasic B, Zeng H. 2021. A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. *Cell* **184**:3222-3241.e26. doi:10.1016/j.cell.2021.04.021

Zebhauser PT, Cordts I, Hengel H, Haslinger B, Lingor P, Akman HO, Haack TB, Deschauer M. 2022. Characterization of cognitive impairment in adult polyglucosan body disease. *J Neurol* **269**:2854–2861. doi:10.1007/s00415-022-10960-z

Zhou X, Stephens M. 2012. Genome-wide efficient mixed-model analysis for association studies. *Nat Genet* **44**:821–824. doi:10.1038/ng.2310

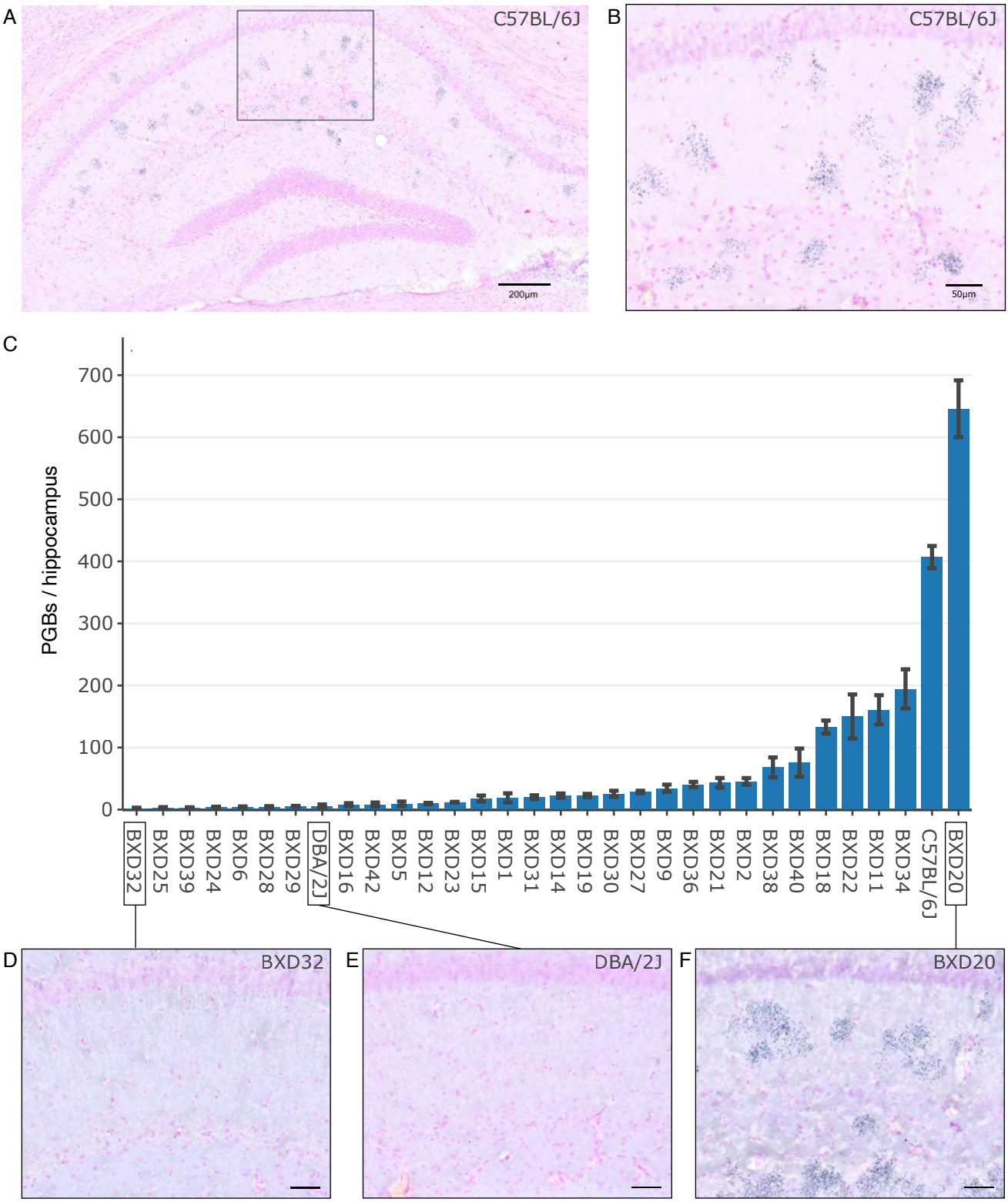


Figure 1. Polyglucosan body accumulation in the hippocampus varies widely between BXD stains. (**A**) LBP110 staining in the hippocampus of an 18 month old C57BL/6J female. (**B**) Zoom in. (**C**) Average of PGBs per hippocampus of 32 BXD strains, as well as the parental strains C57BL/6J and DBA/2J. Genenetwork trait ID: BXD_10685. (**D-F**) Representative pictures of LBP110 staining in the hippocampus of the BXD32, DBA/2J and BXD20 strains. Scale bars represent 50µm.

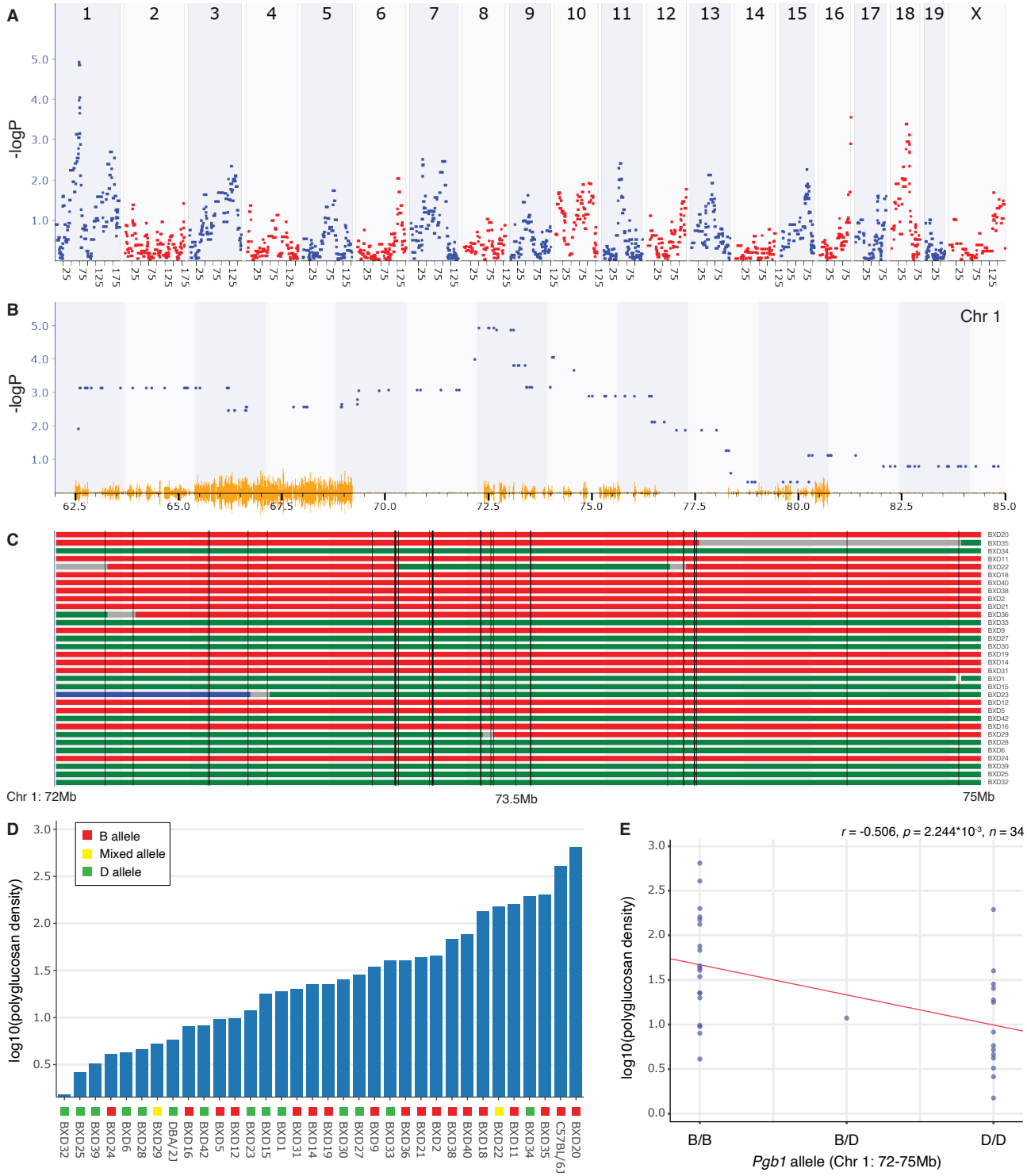
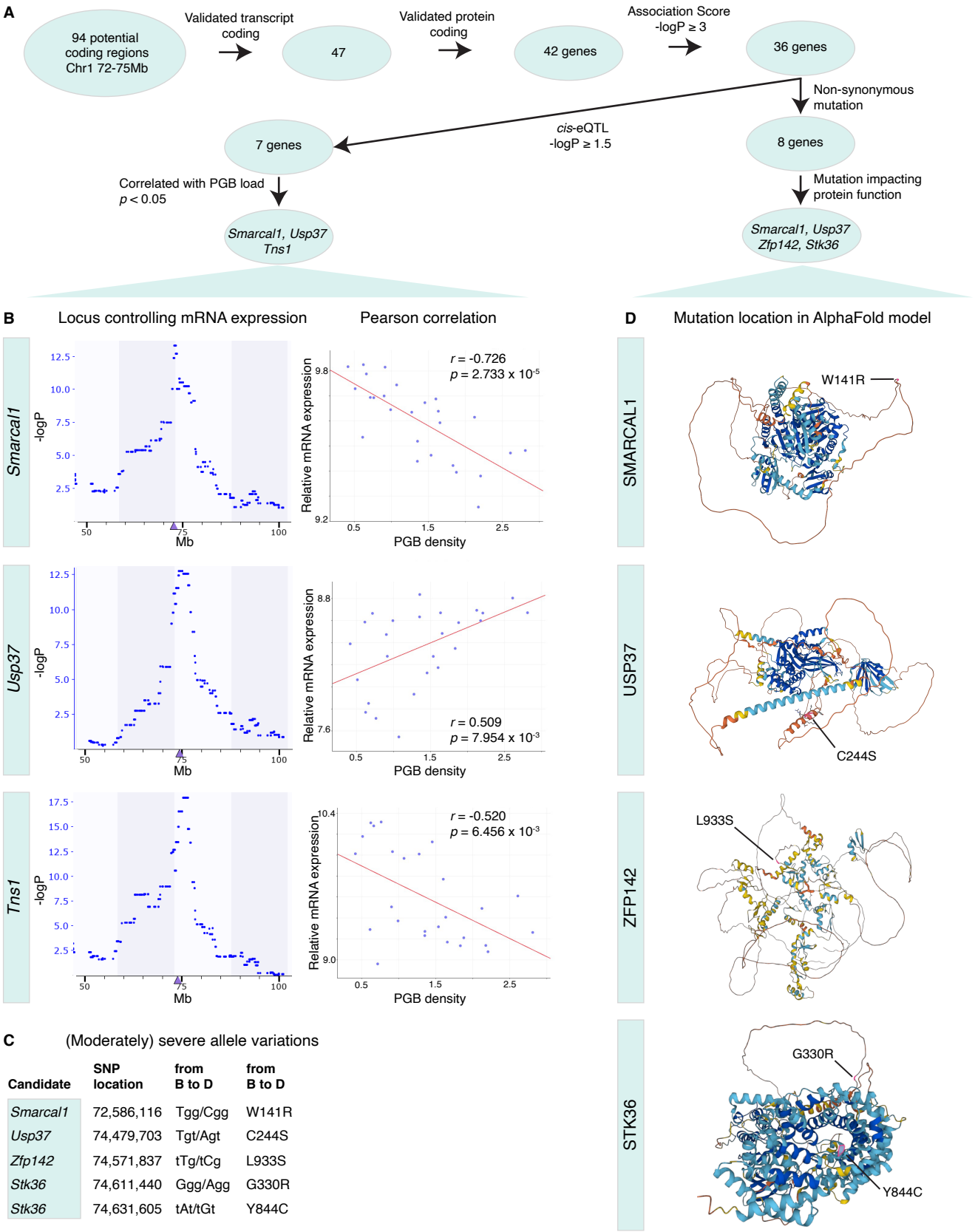


Figure 2. Quantitative trait locus mapping of polyglucosan body density reveals a causal variation on Chromosome 1. (**A**) Manhattan plot indicating the correlation significance between polyglucosan body densities (genenetworks trait ID: BXD_10686) and linkage blocks across the genome. (**B**) Zoom-ins on the target region including the peak region of interest between 72 and 75 megabases (Mb). The yellow color indicates allele variations between BXD strains annotated at genenetworks.org. (**C**) Allele inheritance from either the *B* (red) or *D* (green) parental strain within the *Pgb1* locus. Blue indicates a heterozygous allele. (**D**) Log10 transformed average of PGBs per hippocampus of 32 BXD strains, as well as the parental strains C57BL/6J and DBA/2J. (**E**) Pearson collelation between the *Pgb1* alleles and PGB density.



C (Moderately) severe allele variations

Candidate	SNP location	from B to D	from B to D
Smarcal1	72,586,116	Tgg/Cgg	W141R
Usp37	74,479,703	Tgt/Agt	C244S
Zfp142	74,571,837	tTg/tCg	L933S
Stk36	74,611,440	Ggg/Agg	G330R
Stk36	74,631,605	tAt/tGt	Y844C

Figure 3. Candidate selection on genes located in *Pgb1* reveals candidate variants with a potential impact on PGB density. (**A**) Workflow of candidate selection, which was performed employing the data provided in Table S1. (**B**) *Cis*-eQTL peak location (left) and correlation of candidate gene mRNA expression with polyglucosan body burden (right). n = 26 strains for each pearson correlation plot. (**C**) List of (moderately) severe allele variations between the *B* and *D* alelle within candidate protein coding genes (data extracted from Wang et al. 2016). (**D**) Alphafold models of the candidate genes, produced using the Ensembl browser and genome assembly GRCm39. Indicated are the locations of the allele variations (pink) between the B6 and D2 strains listed in (C). Colors correspond to AphaFold model confidence ranging from very high (dark blue) and confident (cyan) to low (yellow) and very low (orange).

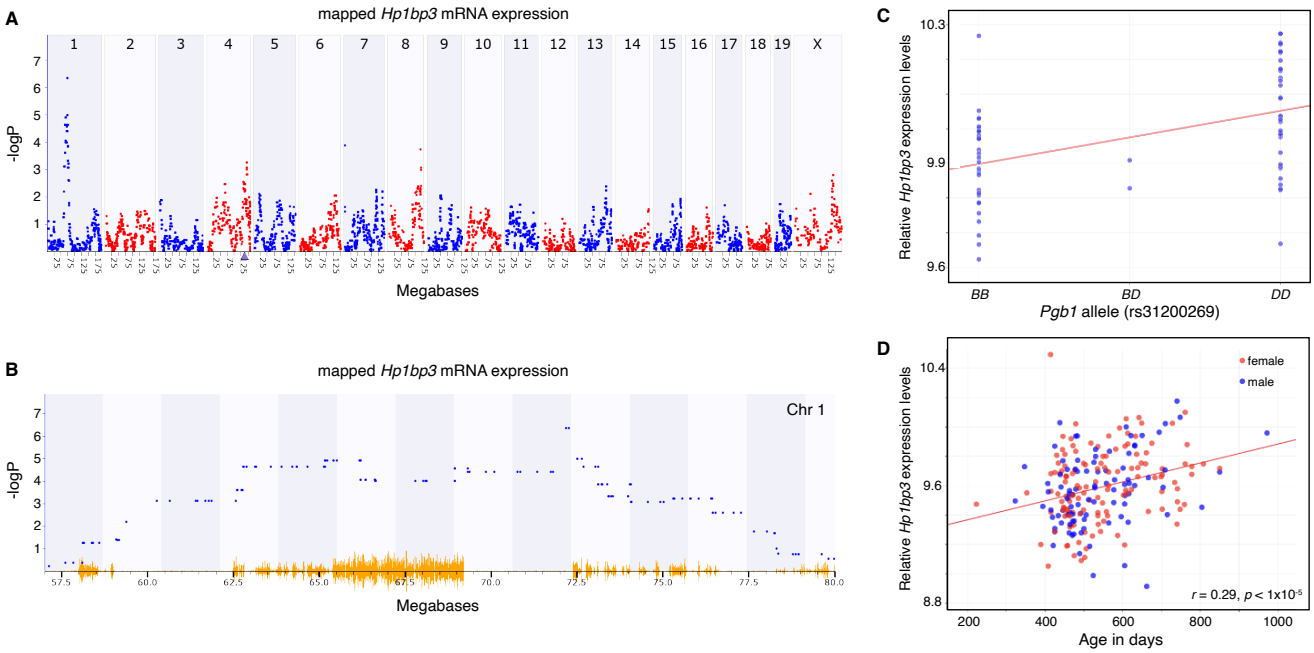


Figure 4. Hp1bp3 is associated with PGB burden. (**A**) The *Hp1bp3* transcript maps as a *trans*-eQTL to the location of the *Pgb1* locus on Chr 1 at ~72.2 Mb. Note that there is a *cis*-eQTL at the location of the cognate *Hp1bp3* gene on Chr 4 at 138.2 Mb (purple triangle on x-axis). (**B**) Zoomed-in view of the trans eQTL. The highest –logP value is ~300 Kb downstream of *Smarcal1* at 72.54 Mb. Yellow indicates allele variations annotated at genenetwork.org. (**C**) Scatterplot of the expression of *Hp1bp3* as a function of *B* and *D* alleles highlighting the protective effect of the *D* allele. (**D**) Scatterplot of the expression of *Hp1bp3* as a function of age. Male (blue) and female (red) cases.

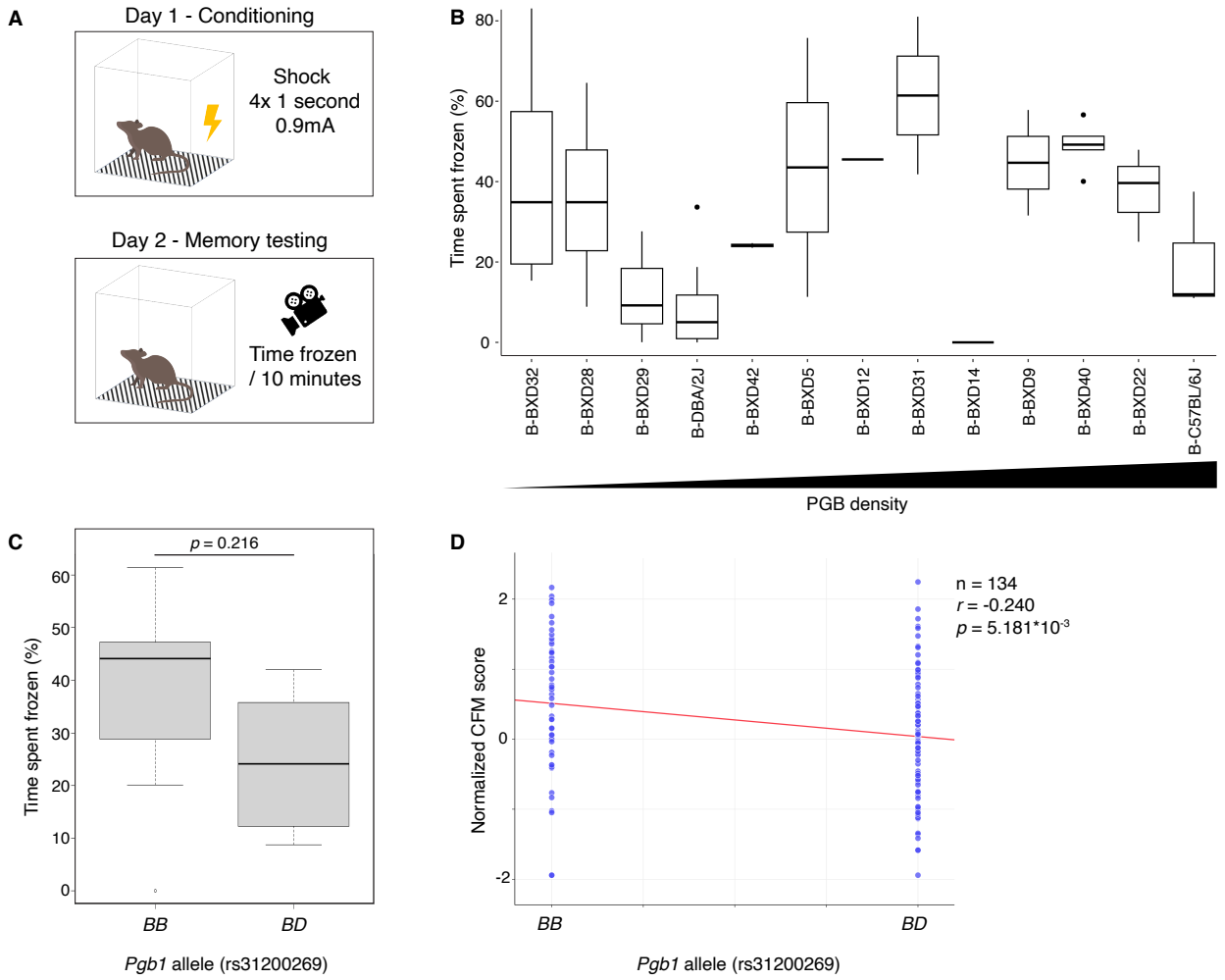


Figure 5. Conditioned fear memory showed no correlation with PGB density in 14-month old mice. (**A**) Workflow of the conditioned fear memory test. (**B**) Time spend frozen plotted for each strain, ordered from low to high PGB density. All tested mice were 14-months of age and resulted from a backcross of homozygous BXD strains with the B parental strain. (**C**) Time spend frozen (%) by *Pgb1* allele. Means per strain (n = 13) were used to account for variable numbers of mice used per strain. No significant correlation was found between the *Pgb1* allele and CFM score. (**D**) Pearson correlation scatterplot showing normalized conditioned fear memory score per allele at *Pgb1* (rs31200269). The correlation can be found under genenetworks ID GN 10474. In total, 134 animals of 14-months of age were tested, across 42 different strains. A significant correlation was found, in which the *Pgb1* (rs31200269) *BB* allele correlates with a higher CFM score.

Table 1A. *Trans-acting regulatory effects of* Pgb1 *on mRNA expression in the hippocampus*

Gene	Location (Mb)	Functional association	–logP	B/D[b]	r[c]	Effect size	Record ID
Hp1bp3	Chr 4: 138.241495	Heterochromatin organization	6.31	D	0.38	0.094	10340721
Pacsin2	Chr 15: 83.375607	Intracellular vesicular transport	3.29	D	0.28	0.050	10430997
Tspan3	Chr 9: 56.131723	Membrane complex stability	3.39	B	0.35	-0.039	10593842
Ext2	Chr 2: 93.695631	Glycosyltransferase active at the Golgi/ER	4.19	B	0.20	0.072	10485225
Eid2	Chr 7: 28.267881	Negative transcriptional regulator	3.19	D	0.33	-0.030	10551483
Cd81	Chr7: 143.052795	Membrane structure and regulation	3.32	B	0.18	0.056	10559261
Fosb	Chr 7: 19.302696	Transcription factor complex	3.16	D	0.57	0.080	10560481
Htr5a	Chr 5: 27.841947	Serotonin receptor	2.55	D	0.47	0.048	10520355
Mtap2	Chr 1: 66.175329	Microtubule cytoskeleton organization	2.68	B	0.45	0.032	10347036

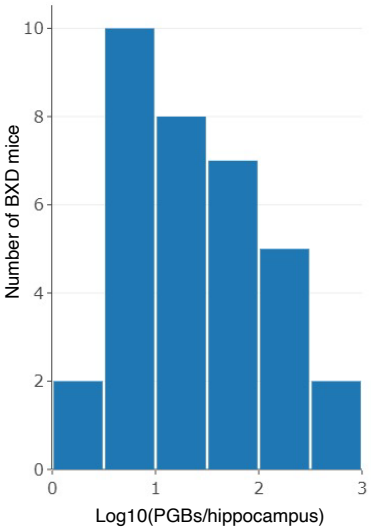
[a]Functional information derived from UniProt and Gene Ontology (UniProt Consortium, 2023; Aleksander et al., 2023). [b]Allele that increases expression. [c]Spearman correlation between PGB burden and mRNA expression.

Table 1B. *Trans-acting regulatory effects of* Pgb1 *on protein expression in the hippocampus*

Protein	Location (Mb)	Functional association[a]	–logP	B/D[b]	Effect Size[c]	Record ID
GSTP2	Chr19: 3.911203	Glutathione tranferase	3.20	B	–0.11	903_ALPGHLKPFETLLSQNQGGK_3
UBA1	Chr X: 20.235452	Ubiquitin conjugation	4.40	B	-0.08	7662_LAGTQPLEVLEAVQR_2
RHOB	Chr 12: 8.497758	Signalling GTPase	3.63	B	–0.10	22979_HFC
ARHGDIA	Chr11: 120.438549	Regulation of Rho GTPases	2.82	B	-0.07	12931_VAVSADPNVPNVIVTR_2
PPP1R7	Chr1: 95.239431	Phosphatase subunit	2.00	B	-0.05	53260_SLETVYLER_2
NPEPPS	Chr11: 97.066948	Aminopeptidase	2.60	B	-0.16	75047_DLSLPPVDR_2

[a]Functional information derived from UniProt and Gene Ontology (UniProt Consortium, 2023; Aleksander et al., 2023). [b]Allele that is associated with increased protein expression. [c]Effect size measured in units of $\log_2$ polyglucosan density per *B* to *D* allele substitution (for example, see figure 4C).

A



B

#	Strain	PGB burden	Animals quantified	#	Strain	PGB burden	Animals quantified
1	BXD32	0.176	3	18	BXD19	1.354	8
2	BXD25	0.415	5	19	BXD30	1.405	4
3	BXD39	0.512	2	20	BXD27	1.453	10
4	BXD24	0.613	5	21	BXD9	1.537	5
5	BXD6	0.623	4	22	BXD33	1.602	3
6	BXD28	0.663	6	23	BXD36	1.607	2
7	BXD29	0.716	5	24	BXD21	1.635	3
8	DBA/2J	0.763	5	25	BXD2	1.657	4
9	BXD16	0.903	6	26	BXD38	1.833	3
10	BXD42	0.916	2	27	BXD40	1.879	3
11	BXD5	0.978	2	28	BXD18	2.124	5
12	BXD12	0.991	3	29	BXD22	2.176	3
13	BXD23	1.072	6	30	BXD11	2.206	4
14	BXD15	1.250	6	31	BXD34	2.289	3
15	BXD1	1.276	3	32	BXD35	2.301	3
16	BXD31	1.299	5	33	C57BL/6J	2.609	4
17	BXD14	1.352	2	34	BXD20	2.810	5

Figure S1. Log transformed polyglucosan count distribution and number of mice investigated for determination of strain-specific polyglucosan body density. (**A**) Polyglucosan accumulation in the hippocampus shows a normal log-transformed distribution across BXD strains. (**B**) The median number of PGBs per hippocampus was determined from an average of four animals per strain.

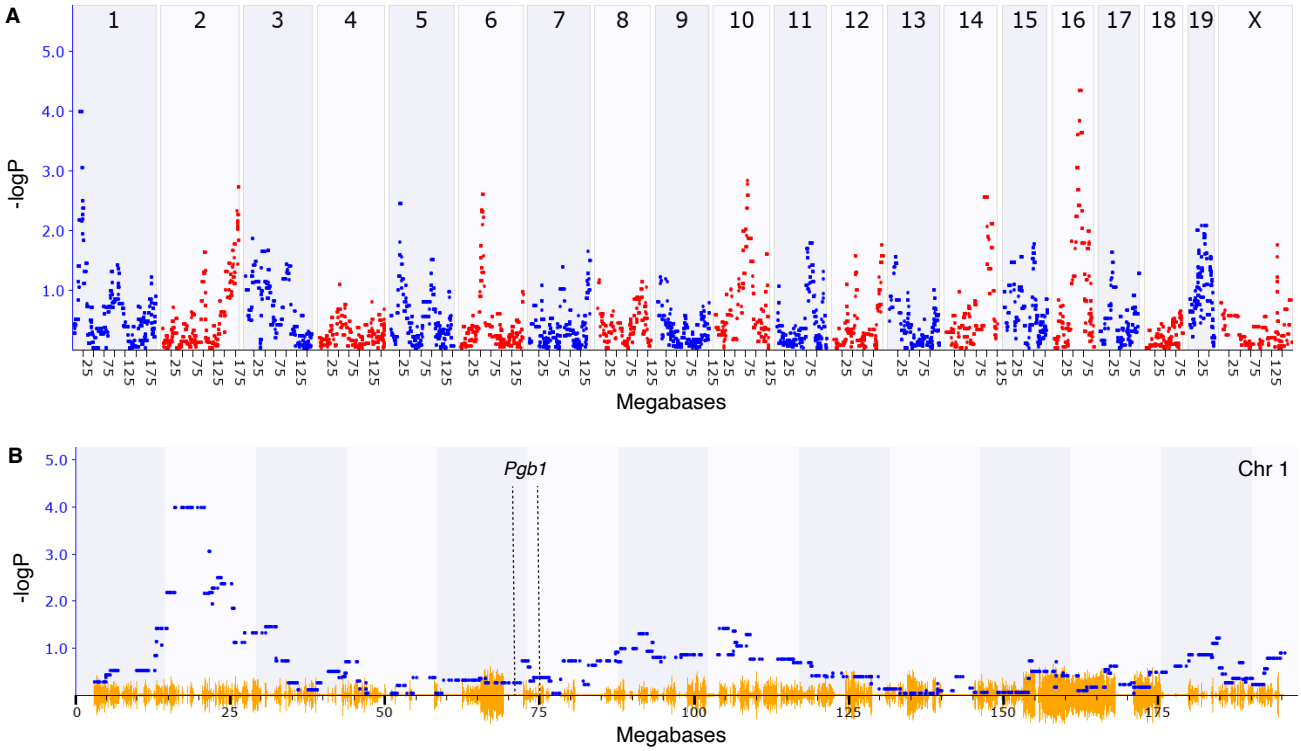


Figure S2. Conditioned fear memory performance does not map to the *Pgb1* locus. (**A**) Mapping of the fear conditioned memory performance using GEMMA shows two major correlating loci (-logP ≥ 4) on chromosome 1 and chromosome 16. (**B**) No correlation was found within the *Pgb1* locus on chromosome 1.