Authors

Lousada E[1,*], Kliesmete Z[2,*], Janjic A[2,*], Burguière E[1], Enard W[2], Schreiweis C[1, §]

Author Affiliations & Addresses

[1]Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, Inserm, CNRS, AP-HP, Hôpital de la Pitié Salpêtrière, Paris, France

[2]Anthropology and Human Genomics, Faculty of Biology, Ludwig-Maximilians University, Munich, Germany.

* These authors contributed equally to the publication

Corresponding authors

[§]Christiane Schreiweis

Email: c.schreiweis@icm-institute.org

Phone: +33-1.57.27.44.29

Title

Expression profiling of the learning striatum

Abstract

During cortico-basal ganglia dependent learning, relevant environmental information is associated with certain outcomes; such learning is essential to generate adaptive behaviour in a continuously changing environment. Through repetitive trial-and-error experiences, actions are optimized and cognitive associative load can be relieved through consolidation and automatization. Although the molecular basis of learning is well studied, region-specific genome-wide expression profiles of the striatum, the major input region of cortico-basal ganglia circuits, during learning are lacking. Here we combined an automated operant conditioning paradigm with an efficient RNA-sequencing protocol to compare expression profiles among three learning stages in three striatal regions per hemisphere in a total of 240 striatal biopsies. Notably, the inclusion of matched yoked controls allowed reliably identifying learning-related expression changes. With 593 differently expressed genes (3.3% of all detected genes), we find the strongest effect of learning at an early, goal-directed stage across all three striatal region and identify a total of 921 learning-related expression changes. Our dataset provides a unique resource to study molecular markers of striatal learning.

Introduction

The way animals learn is a subject of utmost relevance. Learning how to look for food, how to play piano, or how to drive a car requires a multitude of steps until efficiency. The first step requires motivation (for example, hunger, desire to play music, need to drive to work, respectively), and also high amounts of exploration and trial-and-error. After having established the correct associations between environmental stimuli, executed actions and outcomes, performance increases fast. Finally, through sufficient repetition, performance saturates with consolidation and automatization of the acquired task or skill. This important step allows for perfecting our actions and permits freeing the cognitive load required in the previous stages.

Cortico-basal ganglia circuits are crucially implied in task and skill learning and consolidation. They are topographically organized into parallel limbic, associative and sensorimotor loops coursing through the ventromedial (VMS), dorsomedial (DMS) and dorsolateral striatum (DLS), respectively (Hintiryan et al., 2016; McGeorge & Faull, 1989; Pan et al., 2010). These circuits dynamically interact and are recruited to different extents during learning. The limbic cortico-striatal loop, the "reward circuit", is especially important during initial acquisition, when behaviour is highly exploratory and reward-dependent. The associative loop is increasingly recruited during the goal-directed, declarative phase of learning, when creation and strengthening of "good associations" need to be made. Finally, the sensorimotor loop is crucial when the learned behaviour becomes automatized and procedural (Alexander et al., 1986; Graybiel, 2008; Thorn & Graybiel, 2014; White & McDonald, 2002; Yin et al., 2009; Yin & Knowlton, 2006). The striatum forms the major input region of the basal ganglia and receives glutamatergic projections from the cortex and thalamus, which then conveys back into the cortex, through the thalamus. In addition, the striatum receives important innervation and modulation by the dopaminergic midbrain neurons, which are crucially involved in reward-related learning and motivation. Finally, one other important player in the modulation of the striatal circuits in the context of learning is the ensemble of cholinergic interneurons, with the hypothesis of their involvement in the encoding of "gain" of learning and plasticity (Cox & Witten, 2019). Learning leaves enduring neuroplastic traces as a neural engram in involved circuitry. Such neuroplastic changes comprise different phenomena, including for example Hebbian plasticity (Hebb, 1949), post-lesional cortical remapping (Kaas et al., 1983; Merzenich et al., 1983), or myelination (Gibson et al., 2014). To capture comprehensive, dynamic snapshots of the overall learning engram systematically and to identify learning-affected procedures and genes in a data-driven manner, we designed an automated, self-motivated learning task. Our experimental design allowed us to characterize different learning stages systematically and to generate a sufficiently high number of samples to set out to develop a molecular map of the learning striatum. Importantly, sets of control animals allowed us to identify changes that stem from learning. Concretely, we developed custom-modified operant conditioning chambers (Benzina et al., 2021), which, in contrast to standard laboratory operating procedures, serve both as housing as well as learning environment for the animals. Such design allows for conditioning the animals

in a self-paced, self-motivated manner and with reducing experimental stress to a possible minimum. Sets of such chambers allowed for generating cohorts of animals for early, intermediate and expert learning stages, hereby enabling the collection of a high number of samples and reducing noise through controlling for cohort effects. Finally, each learning mouse was paired with a matched control animal in a mock chamber in the same experimental room, controlling for experimental side effects other than learning-related effects. Additional naïve mice, which never entered any treatment, served as an outgroup control for any experimental procedures. Sample harvesting for genetic expression analyses was performed at precisely defined time points from three different striatal regions, namely limbic, associative and sensorimotor striatum, which have been all described as relevant to a different extent for different learning stages. We clustered observed genetic expression changes according to their dynamics across learning and identified annotated gene categories implied across learning in general as well as in particular to each learning stage in a first draft of a comprehensive molecular map of the learning striatum.

Materials and Methods

Animals

All experimental procedures followed national and European guidelines and were approved by the institutional review boards (French Ministry of Higher Education, Research and Innovation; APAFiS Protocols no. 1418-2015120217347265 and 2021042017105235). Animals were group-housed in the animal facilities of the Paris Brain Institute in Tecniplast ventilated polycarbonate cages under positive pressure with hard-wood bedding in groups of up to six animals per cage with *ab libitum* food and water access. The temperature was maintained at 21–23 °C and the relative humidity at 55 ± 10% with a 12-h light/dark cycle (lights on/off at 8am and 8pm, respectively). The animals were maintained in a 12-hour light/dark cycle (lights on/off at 8:00am/8:00 pm, respectively), and had *ab libitum* food and water access. Following ethical guidelines of animal experimentation on gender equity and the principle of the three Rs, we included both male (n = 99) and female adult mice (n = 60) (aged 5.50 ± 1.88 months). For behavioural task validation, wildtype animals (n = 25, hereof n = 19 males and n = 6 females) originated from several mouse lines on C57BL6 background in order to reduce breeding for ethical reasons.

Experimental setup and task

Operant conditioning was conducted in custom-modified experimental chambers (ENV-007CTX, Med Associates, Vermont, USA) as previously described (Benzina et al., 2021), in which each of the tested animals lives and performs the conditioning task 24/7 in an automated, self-initiated, self-paced manner. Seven such experimental operant chambers were operated in parallel in the same experimental room. Briefly, the rear wall of each chamber housed the feeder compartment (equipped with an infrared light beam to detect feeder entry crossing) and drinking water bottle holder, the front wall held two tactile screens, and a custom-made pair of black Plexiglas gates equipped with a pair of infrared (IR) beams.

These beams allowed for tracking the locomotion of the mouse into or out of the area with the tactile screens. The chambers of the yoked controls were custom-built, plexiglass chambers as described in (Lamothe et al., 2021) and were located in the same experimental room as the operant conditioning chambers. The yoked control chambers contained ad libitum water access as well as the same woodchip bedding and cotton pad nesting material as the operant conditioning chambers and regular housing cages in the facilities. Precision reward pellets served as the sole nutrition during the entire experiment. The reward pellets (20 mg LabTabTM AIN-76A rodent precision pellets, TestDiet, Richmond, USA) are earned upon successful trials by the mice in the conditioning task ("learners"); yoked control mice, matched for sex, age and whenever possible also for litter, receive the exact same amount of food rewards as their matched pair undergoing actual conditioning.

Once placed in the operant conditioning chambers, the mice start with a pre-training phase of four different stages. In a first stage, a total of five precision pellets drops automatically once every minute if the previous pellet had successfully been recovered. In a second stage, the mouse is required to nose poke into the feeder compartment to trigger individual pellet delivery once per minute until a total of 10 pellets has been received. In a third stage, crossing of the IR photobeams prompts the illumination of both tactile screens; the mouse is required to touch one of them to trigger pellet delivery in the feeder. There was no time limit for recovering the pellets at this stage. However, the condition for the mice to pass into the following pre-training stage was the recovery of the pellets within less than 15 seconds for at least 66% of the times. Then, in the consecutive pre-training stage, as before, both screens were illuminated with white light upon gate crossing. However, mice now had to nose poke into the feeder in less than 15 seconds after screen touch in order to trigger pellet delivery. Having successfully recovered at least 66% of the pellets in this manner, mice passed into the actual task.

The behavioural task was a deterministic, binary visual discrimination task, where mice self-initiated each trial by triggering visual stimulus onset on the tactile screens by again crossing the photobeam-equipped gate. The presented stimulus pair consist in ten, either vertical or horizontal, white bars on a black screen, each of which was presented 50% of the time in a random manner. To earn a precision pellet, mice need to associate each of the two stimuli with one screen touch, e.g. stimulus of ten horizontal/ vertical bars associated with the action of left/right screen touch, accordingly. The contingencies were balanced across individuals and operant chambers. Mice were allowed a maximum of 60 seconds after stimulus onset to respond with a screen touch. Upon correct response, the screens started blinking to give a positive feedback for the mouse, which had 15 seconds to nose poke in the food dispenser in the rear wall for obtaining a food reward. Past 15 seconds after correct response, the trial became unrewarded. In case of incorrect screen touch, mice received an aversive white light for five seconds. Additionally, in order to decrease the tendency of mice to lateralize, each trial following an incorrect response required the mouse to switch to the other screen. Past 60 seconds in the chamber compartment with the screens, the screens would switch off again; in these time out trials, mice then had to cross out and back into the feeder compartment to trigger the next trial. Trials where mice crossed

in and out of the screen compartment within 60 seconds and without screen touch, were quantified aborted trials (Supplementary Figure 1).

Animals for validation of the task and subsequent determination of the time points of early, intermediate and late stages of learning remained in the task for around 5500 trials (trial number = 5517 +/- 1177; median +/- SD, respectively). The time point of early learning stage was estimated through averaging individual subjects' inflection points of ramping performance towards the beginning of the performance plateau (trial number = 1240 ± 409; median +/- SD, respectively) (Supplementary Figure 2). Late learning was determined as 2000 trials post early learning stage time point. Finally, intermediate learning stage for each individual was determined as the intermediate time point between individual early and late learning stage time points.

Animals for subsequent RNA-seq were divided into early, intermediate and late learner groups; the groups were balanced for age and litter, and the animals sacrificed the morning after having reached each stage defined at 1000/2000/3000 trials for early/intermediate/late learners, respectively. An additional group of naïve animals, which was never exposed to any behavioural task, served as a control for experimental treatment. To control for experimental factors, each learning mouse was paired with a matched animal of the same genotype, sex, and age as a yoked control (for e.g. change of the experimental room, social isolation, food consumption and quality). Learners and yoked controls were put into their according behavioural setup at the same time, and sacrificed at the same time, with the order of the dissection/tissue harvesting being randomized. The yoked controls received the exact same amount of precision food pellets as its learning partners during the past 24 hours in the mornings.

Sample collection

Having reached respective criteria for early/intermediate/late learning stages, the brains of both learners and their paired yoked partners were rapidly dissected and collected during the morning of the following day. Naïve animals were directly taken from the group-housed home cages and their dissections were done at the same hours as the mice of the other groups.

Before each dissection, all tools and surfaces were cleaned with RNAse decontaminating solution (Fisher Scientific, reference #10180601). Concretely, animals were euthanized by cervical dislocation, rapidly dissected and their brains immediately collected and sliced (1mm thickness) using a 5-blade custom-made tool to rapidly obtain coronal slices of 1mm thickness at comparable levels across mice. All dissections were performed in a petri dish, which had been cleaned with RNAse decontaminating solution beforehand. Brains were sprinkled with sterile saline to remove excess blood. All samples were collected by the same experienced researcher using $1mm^2$ biopsy punchers (pfm medical, reference #49101) to collect $1mm^3$ samples from each hemisphere of the three regions of interest: ventromedial striatum (AP = 0.10/ML = ±1.00/DV = 4.00), dorsomedial striatum (AP = 0.14/ML = ±1.50/DV = 2.10), and dorsolateral striatum (AP = 0.14/ML = ±2.50/DV = 2.50). Samples were rapidly placed inside of a RNAse-free Eppendorf tube (Dominique Dutscher, reference #033872) containing 200µL of buffer RLT

Plus (Qiagen, reference #1053393), and immediately frozen on dry ice. All samples were transferred and kept at -80°C until shipment, which was done on dry ice.
Samples were transferred to 96-well plates and RNA purification and RNA-sequencing was performed with 50µl lysate using the prime-seq protocol as described (Janjic et al., 2022). Separate libraries were generated for the three striatal regions and sequenced together on four lanes of an Illumina NextSeq 2000 P2 flow cell with Read 1 (barcode & UMI): 28 bp, Index 1 (i7): 8 bp; Read 2 (cDNA): 101bp.

Data analysis

Behavioural analyses

All behavioural data were analyzed using Matlab R2017b and the open source package 'estimation stats - DABEST' for data visualization (Ho et al., 2019). For behavioural strategy analyses, we additionally applied the PsyTrack algorithm (Roy et al., 2021). Statistical analysis were performed using IBM SPSS Statistics 29. We used a generalized linear mixed model to investigate the effects on three different targets: performance, number of trials, and pellet recovery times. The model used was target ~ 1 + Age_start + Sex + Mouse_line + Cage + Batch + Learning_stage + (1|Mouse).

Expression data processing and mapping

Fastq files were trimmed to remove any bases from the poly(A) tail using Cutadapt v1.12 (Martin, 2011). The quality was assessed using fastqc v.0.11.8 (ANDREWS, n.d.). zUMIs pipeline v.2.9.4d (Parekh et al., 2018) was used to process the data and generate a count matrix. Reads with a Phred quality score threshold of 20 for 3 BC bases and 4 UMI bases were filtered, mapped to the mouse genome (GRCm38) with the Gencode annotation (vM25) using STAR v.2.7.3a (Dobin & Gingeras, 2015), and then counted using RSubread v.1.32.4 (Liao et al., 2019). Genes that were expressed in at least 20% of the cells with an average of 3 counts or higher across samples from at least 1 striatal region were kept for the downstream analyses. This resulted in a count matrix containing 17,499 genes across 240 samples.

Cell-type composition

SCDC v.0.0.0.9000 (Dong et al., 2021) was used for cell-type deconvolution, using the scRNA-seq TM_Brain_Non_Myeloid reference from Tabula Muris (Schaum et al., 2018). We observed varying proportions of inhibitory neurons (0.6 - 0.88) across samples, however these proportions are likely variable due to technical reasons, not biological, as we did not observe any systematic difference in overall cell-type composition between mouse samples under learning conditions vs controls (multinomial test p-value 1.0).

Model selection and exploratory analyses DESeq2 v.1.34.0 (Love et al., 2014) was used for the initial data exploration and backward model selection (Likelihood ratio test, Benjamini-Hochberg adjusted p-

value <0.05). A full model including all relevant biological and technical factors was fitted using the following design ~0 + Condition + Region + Batch + Hemisphere + prop_inhibitory + Age_start_months. Predictor importance was tested using leave-one-out approach. All predictors besides Age_start_months showed significant improvement in explaining the variance of >200 genes. Age_start_months only explained 2 genes and was further removed from the model. Another round of leave-one-out model selection was performed without including Age_start_months, but none of the kept predictors showed reduced importance, hence all other predictors were kept. Principal component analysis (PCA) and hierarchical clustering (Euclidian distance, method "complete") using R package pheatmap v.1.0.12 (Kolde, 2019) were performed on the 1000 most variable genes (centered and scaled matrix), and we observed clustering mainly due to the striatal region, but also cell-type composition, batch and learning effects.

Differential expression analysis

Input expression matrix was normalized using scaling factors calculated via TMM method from R package edgeR v.3.36.0 (Robinson & Oshlack, 2010). Differential expression (DE) analysis was performed using DREAM (Hoffman & Roussos, 2021), R package variancePartition v.1.24.1 (Hoffman & Schadt, 2016), which implements mixed-effect modelling. Mouse pairs were included as a random effect in addition to the main predictors specified above, resulting in the following final model: 0 + Condition + Region + Hemisphere + Proportion (IN) + Batch + (1|Mouse pairs). The importance of the included predictors was validated using variance partition implemented in DREAM. Benjamini-Hochberg adjusted p-value < 0.05 was considered as a cut-off for significant DE.

Clustering of differentially expressed genes

The included genes consisted of genes that were DE between either learning stage (GD, Int, Hb) and its respective control (yokedGD, yokedInt, yokedHb), respectively, or the combined (averaged) contrast across the learning stages versus controls, comprising a total of 921 learning-associated genes. The log2-fold changes to the respective controls were used to cluster these genes using k-means unsupervised clustering algorithm (Eucledian distance on centered and scaled matrix), R package factoextra v.1.0.7 (Kassambara & Mundt, 2020). The relevant number of clusters was inferred using three cluster metrics: silhouette, elbow and gap statistic. All of these preferred cluster number (k) of 2-4, showing very similar preference among these 3 options. Hence, we chose to evaluate the clusters using the range of 2-4 k.

Gene set enrichment analysis (GSEA)

GSEA was performed using R package topGO v.2.46.0 (Alexa & Rahnenfuhrer, 2022) Biological Process (BP), setting minimal node size to 10. Significantly enriched terms (Fisher's p-value < 0.01) were further filtered for terms with <500 annotated genes and >2 significant genes to facilitate interpretability.

Software

All expression analyses besides pre-processing were performed using statistical software R v.4.1.3 (R Foundation for Statistical Computing, Vienna, Austria., n.d.). Figures and data wrangling were done using tidyverse v.1.3.2 (Wickham et al., 2019) and cowplot v.1.1.1 (Wilke, 2020).

Results

A deterministic visual discrimination task to assess learning

To capture learning in the most naturalistic possible manner, we have developed a fully automated, self-motivated, visual discrimination task. Mice live inside custom-modified operant conditioning chambers and self-initiate each trial in order to obtain food rewards by crossing an infrared beam-equipped gate (Fig. 1A). By doing so, the two touchscreens inside the chamber wall facing the mouse display simultaneously either ten vertical or horizontal bars as a stimulus. Each mouse was assigned to a certain stimulus-outcome contingency, e.g. horizontal/vertical bars require a left/right screen touch, respectively. We evaluated all trials during the animals' awake cycle (8pm-8am) and distinguished between five response types: correct and incorrect trials, aborted trials, time out trials and trials, in which the food reward was not picked up within 15 seconds after correct response (Supp. Fig. 1). Wildtype mice acquired the task, following the expected learning curve with a steady increase in initial performance and plateauing performance at trial number = 1240 ± 409 (Fig. 1B, Suppl. Fig. 2). We thus distinguished four different learning stages: naïve, early, intermediate, and late. The naïve stage corresponds to the first night in the chamber. The early learning stage was defined as the night in which the performance plateau was reached (Suppl. Fig. 2). Subsequently, intermediate and late learning stages were defined as 1000 or 2000 trials past early learning, respectively (Fig.1B). Performance levels differed between early and both intermediate and late learning stages, while this was not the case between the two latter stages. Besides performance, we used additional behaviourally relevant metrics to further characterize the different learning stages: reaction time, i.e., the time between stimulus onset and screen touch; pellet recovery time, i.e., the time between correct screen touch and nose poke to trigger pellet delivery; and the number of trials, which we additionally segregated into five response types. Reaction time was highly variable between animals and did not reflect modulation across learning (Supp. Fig. 3). Pellet recovery time steeply decreased during initial learning before plateauing (Fig. 1C). The overall number of initiated trials decreased from early to subsequent learning stages. Indeed, subtracting performance and initiated trials at early learning stage from those of a later learning stage highlights an increase in efficiency across learning (Fig. 1D). We performed generalized linear mixed model analysis on the response variables "performance", "number of trials", and "pellet recovery time", using learning stage, age, sex, mouse line, cage, and batch as predictors, and correcting for animal as random effect. For performance and pellet recovery time, the only significant predictor was learning stage (Performance: F: 152.976, df1: 3, df2: 78, p-value<0.001; Pellet recovery time: F: 66.554, df1: 3,

df2: 78, p-value<0.001). As for number of trials, no fixed effect reached statistical significance, being however close (Number of trials: F: 2.536, df1: 3, df2: 78, p-value=0.063).

While performance and rapidity are unquestionably important parameters to assess learning over time, we believe that understanding the hidden strategies that drive a subject to take a decision can substantially provide insights into behavioural optimisation. To analyse such hidden strategies, we used the PsyTrack model (Roy et al., 2021). Such model allows to estimate the weights of different strategies that can be applied by the animals. Concretely, in our task, we identified four relevant strategies: in the "current contingency" strategy, the animal follows the assigned stimulus-outcome contingencies and represents the optimal strategy for this task. The "lateralization bias" reflects the natural tendency of each mouse to choose one side over the other. The "previous contingency" strategy represents win-stay/loose-shift behaviour. And lastly, animals could also apply a simple repetition/alteration strategy. In our task, we observed an increase of the weight of following the assigned contingencies across learning, indicating that the animals correctly apply the acquired contingencies as a dominant task strategy (Fig. 1E). Such increase closely follows the performance curve, suggesting that the increase in performance mainly is due to understanding the task structure. Lateralization bias, the initially most dominant strategy, and repetition/alternation strategy remain stable over time. Interestingly, the contribution of the previous contingency strategy slightly increases during initial learning, and completely extinguishes near the time that the current contingency strategy reaches its plateau. In summary, different metrics extracted from the developed task suggest learning as well as learning-related changes across time. Such difference is especially visible between naïve, early and intermediate learning stages, but not beyond.

Replication of task metrics in mice used for RNA-sequencing

Having validated the assessment of learning and behavioural changes across the different learning stages in a first cohort of wildtype mice, we now aimed to track molecular changes across learning and consolidation. To do so, we repeated the behavioural task in a separate cohort of wildtype mice (n = 40), which were assigned to different learning (naïve, early, intermediate and late) and control groups (yoked groups for early, intermediate and late learner groups) (n = 6 per group except the naïve group with n = 4) (Fig. 2A). Group assignment was balanced for age and home cage. In a slight deviation from the original task, we pre-determined a common time point for the early learning stage at 1000 trials instead of attributing an individual time point to each subject in order to extract information from individual learning variability. Subsequent intermediate and late learning stages in this second cohort were thus also uniformly defined at 2000 and 3000 trials, respectively. Having reached the required number of trials, each group was sacrificed the morning after and bilateral striatal punches from the ventral, dorsomedial and dorsolateral striatum taken for subsequent RNA-sequencing procedures. In order to control for experimental condition, handling and food intake, each learner mouse of the early, intermediate and late learning stages was paired with a yoked control mouse. This yoked control mouse

would stay in a mock chamber in the same experimental room for the same amount of time, receiving exactly the same amount of food pellets as its paired learner mouse. The yoked animals were dissected at the time as the learners, hereby randomizing the dissection order across yoked and learner groups. Naïve animals served as an external control for any experimental condition.

We first confirmed that the behavioural parameters obtained in the second cohort was comparable to the ones obtained from the first wildtype cohort. Indeed, as in the first wildtype cohort, performance steeply increased from the early to intermediate learning stages, but not between intermediate to late learning stages (Fig. 2B and 2C). Similarly, the total number of trials decreased between early and the subsequent learning stages, yet without any further decrease between intermediate to late learning stages (Fig. 2D). Given reduced animal number (n = 6 per group instead of n = 25 in the first cohort), the decrease in the pellet recovery times was only observed as a tendency (Fig. 2E). We then correlated the number of trials initiated by the mouse with the performance achieved during the same night and we can notice a clear separate cluster for early time point that shows worst performance for a bigger number of initiated trials. Finally, we conducted generalized linear mixed model analysis on the response variables "performance", "number of trials", and "pellet recovery time", using learning stage, age, cage, and batch as predictors, and correcting for animal as random effect. Both performance and number of trials were strongly predicted by learning stage, with additionally significant effect of batch (Performance: Learning stage – F: 6.781, df1: 2, df2: 6, p-value=0.029; Batch – F: 6.375, df1: 3, df2: 6, p-value=0.027; Number of trials: Learning stage – F: 11.223, df1: 2, df2: 6, p-value=0.009; Batch – F: 8.646, df1: 3, df2: 6, p-value=0.013). Pellet recovery time was not predicted by any of the effects.

RNA-sequencing analysis shows learning-related differential gene expression

We processed the 240 striatal punch biopsies (40 mice x 3 regions x two hemispheres) using the prime-seq protocol as described (Janjic, et al, 2022). In this plate-based bulk RNA-seq protocol, RNA is purified directly from the RLT dissolved biopsies using magnetic beads and subsequently reverse transcribed with barcoded oligo-dT primers. This allows pooling of barcoded cDNA before second strand synthesis, amplification and library generation and is an efficient and sensitive way to generate many libraries (Janjic, et al, 2022). We generated on average 5.3 million reads per sample, resulting in ~ 3 million reads mapped to genes of which 1.5 million were unique reads (UMIs). To remove noise, we selected the 17,499 genes that are in at least one region expressed in at least 20% of the samples with an average of at least three UMIs. We deconvoluted the expression profiles using scRNA-seq data from annotated brain cell types to analyse cell type compositions. As expected, the great majority of cells in our samples were inhibitory (60-88%), followed by excitatory, and glial cells (Fig. 3B). Importantly, the proportion of inhibitory cells was not only a good predictor for variations in other major cell types, but also explained much of the variance observed in the principal component analysis (PC 2, 18% variance, Suppl. Fig 4). Given the relevance of this variability, and to avoid excluding potential outlier samples, we included the proportion of inhibitory cells as a factor in our statistical model. Through backward

model selection further important covariates such as hemisphere, region and batch were included as fixed effects (see methods section for details). Given the experimental design of learner and paired yoked animals, we defined mouse pair as the random effects predictor. We used a mixed-effects differential expression analysis framework (DREAM, (Hoffman & Roussos, 2021)) to infer differentially expressed (DE) genes due to learning, while accounting for both random and fixed effects in our data . We observe that a large proportion of variation in gene expression can be accounted to regions. Firstly, hierarchical clustering of the 1000 most variable genes in our dataset indicates that variation mainly results from regional punching (Fig. 3C). Also, when analysing what percentage of gene expression variance can be explained by either predictor, the largest differences across the data set are between regions (Fig. 3D), however condition still explains 5-29% of variation in 557 genes. Interestingly, also two groups of samples seem to cluster according to learning/yoked condition irrespective of the region (Fig. 3C). Furthermore, these two clusters come almost exclusively from dorsomedial and dorsolateral, but not from ventral striatal samples.

Our experimental design allowed for a multitude of comparisons and analyses (Fig. 3E). We first started by comparing naïve and yoked control groups to detect expression changes, which were attributable to experimental testing other than learning. Given the near absence of gene expression differences between naïve and control groups, we excluded major side effects due to experimental condition per se. Importantly, this meant that the expression differences between yoked and learning groups could be interpreted as learning-related. When segregating those learning-related differential gene expression into the three learning stages, we detected substantial expression changes during early learning , indicating a peak of intense molecular activity by 593 differentially expressed genes (for all mentioned DE results, adjusted p-value<0.05). Past the early learning stage, differential gene expression tremendously decreases from nearly 600 to 55 differentially expressed genes between intermediate learners and yoked controls, to only 36 differentially expressed genes between late learners and yoked controls. Next, we contrasted differential gene expression between learning stages for an imprint of striatal learning across time and detected similar amounts of differentially expressed genes between early/intermediate, and intermediate/late comparisons. Highest differential gene expression was detected between early and late learning stages. We next assessed the communalities and particularities in gene expression patterns between the learning stages (Fig. 3F). As of no surprise, the majority of differentially expressed genes were specific to the early learning (327 genes), whereas there were very few differentially expressed genes specific to the intermediate or late stages (14 and 12, respectively). To add functional information, we selected all learning-related differentially expressed genes (921 genes in total) and annotated gene functions using gene set enrichment analysis of gene ontology “Biological process” (Fisher’s p-value<0.01, (Alexa & Rahnenfuhrer, 2022)) (Fig. 3G). Interestingly, many synapse-related categories appeared in this analysis, including “regulation of long-term synaptic potentiation”.

Comparison between consistently and early expressed genes

We clustered all 921 learning-modulated genes across subsequent learning stages using three cluster metrics: silhouette, elbow and gap statistic (Supp. Fig 4). All of these preferred cluster number (k) of 2-4, showing very similar preference among these three options. Hence, we chose to evaluate the clusters using the range of 2-4 k. With 2 k-means clusters, the separation was mainly between the up- or down-regulated genes, whereas with 3 clusters the up-regulated genes were further subdivided. When dividing these patterns into 4 k-means clusters (Fig. 4), this allowed us to identify 71 consistently up-regulated genes (cluster 1) and 78 consistently downregulated genes (cluster 3). Clusters 2 and 4 (332 and 440 genes, respectively) contained mainly early learning-modulated genes. Consistently upregulated genes were mostly related to angiogenesis, while the consistently downregulated genes were mostly related to circadian regulation. The clusters of mainly early-learning modulated genes contained synapse-related categories, annotations of developmental growth, learning, and others.

Discussion

In the present study, we systematically characterized learning behaviour across defined early, intermediate and expert learning stages in a large number of animals and mapped out learning-related molecular changes. Automated, self-paced, self-motivated operant conditioning as well as a set of matched control conditions allowed us to minimize molecular noise generated by the experimental procedure per se. Operant conditioning set-ups have been used and developed since their first creation in the 19th century (Calkins, 1896; Skinner, 1963; Thorndike, 1927). Despite varying designs and complexity of operant learning paradigms, the vast majority of experimental work still relies on daily conditioning sessions of rather short duration. Such design often includes drastic experimental interferences to elevate the motivational drive of experimental animals to engage in the task such as e.g., water- or food-restriction in mammals (Balcombe et al., 2004). Here, we experimentally minimized such experimental interference and statistically confirmed our approach first in replicable behavioural datasets according to expectations and, second, through molecular expression analyses.

We opted for a simple deterministic binary discrimination task and observed the expected sigmoid performance changes across learning stages with steep performance increase during initial learning followed by plateauing performance towards expert performance. Expert learning never translated into performance at 100%, suggesting the maintenance of an exploratory component throughout the entire task. Similarly, also pellet recovery times, an indicator of motivational association of action and outcome, suggest a dynamic development of learning stages from naïve through early, intermediate towards late learning stages. Indeed, pellet recovery time experiences a sharp drop-off from naïve to early learning, supporting the strongest associative processes between naïve to early learning as observed in the performance parameter. However, the observation of a plateauing effect in isolated yet meaningful behavioural readouts does not indicate static prolongation of early learning states. Indeed, when applying hidden strategy analysis (Roy et al., 2021), opposing dynamics of the strategies following either stimulus-outcome or win-stay/loose-shift strategies favor continued learning beyond the early

learning stage. As an alternative behavioural parameter to assess learning dynamics over time, efficiency serves as an additional readout for dynamic learning changes. Indeed, when contrasting early and late performance over early and late trial numbers, we observed a shift of the majority of the animals into the quadrant of higher performance over less number of trials, indicating an increase in efficiency over time in our task. We did not observe a change of reaction times across learning, which had been observed previously when comparing initial and late learning stages (Thorn & Graybiel, 2014). While this seems surprising at first sight, such observation is likely due to the task design: in tasks such as the T-maze tasks, a widely applied task to study learning-stage related questions, the stimulus-action and action-outcome components are at least partially overlapping while they are completely separated in our task design.

Having observed dynamic changes across learning stages, we designed an experimental protocol suitable to assess a molecular imprint of learning and of learning across the different stages without major interference of experimental factors. Indeed, the differences in expression changes observed between learning animals and paired yoked control groups could not be explained by experimental treatment per se, for which we had controlled for by comparing naïve and yoked control animals, which yielded hardly any expression changes. Similarly, also between-learning stage differences could not be explained by experimental manipulation given the near absence of expression changes between the different yoked control groups. Thus, observed expression changes could be attributed to learning-associated procedures, which were comparable to the first cohort of mice given similar performance, efficiency, applied strategies, pellet recovery times and reaction times. Our dataset could thus be applied to study learning-related questions in general as well as questions addressed to particular subsets of learning stages. Such approach could be informative when studying certain gene categories, networks or single genes of interest that for example have emerged from the literature of neuropsychiatric animal models and that have been suspected to affect associative learning in one way or the other.

Consistent with the dynamics of behavioural parameters, we observed major changes in differential gene expression during early learning, with a decrease in differential expression changes at later stages of learning. We first grouped all learning-related expression changes and analyzed major expression changes across all learning stages in a data-driven manner. The top fifteen striatal learning-regulated gene ontology categories affect three major groups: behavioural categories, such as locomotion or regulation of heart rate; cellular categories such as cell migration; or, to the largest extent, categories describing neuronal molecular changes, in particular neuroplasticity- or general synapse-associated categories including e.g. membrane depolarization, regulation of long-term synaptic plasticity, receptor internalization, regulation of glutamatergic synaptic transmission or acetylcholine receptor signaling. We consider all categories that emerged from such unbiased analyses meaningful in the context of learning and when describing plasticity-related changes in cortico-basal ganglia circuits. Obtaining such meaningful categories further corroborates our approach and results of learning-related differential gene expression.

When pooling and clustering all learning-related genes across segregated learning stages, we obtained four clusters: genes, which were characterized either by a peak of differential gene expression (up- or downregulation) in particular during the early learning phase, or by a consistent up- or downregulation pattern across learning stages. Interestingly, even in more consistently learning-regulated genes, a peak of differential gene expression is present during early learning; this again corroborates the paralleling of behavioural and molecular findings.

When analyzing gene expression categories in a data-driven manner, the top five categories of more consistently upregulated genes across learning stages suggest major upregulation of angiogenetic and cardiovascular processes. This corroborates on the one hand functional hyperemia as a metabolic marker, which forms the principle pf functional MRI BOLD signaling: here, focal blood influx is used to assess neural activity levels given the correlation of both parameters. Beyond such metabolically inspired interpretations of increased blood flow to provide the necessary gas exchange at metabolically highly active sites, one might also reconsider our observations in the context of hemodynamic neuromodulatory functions (Moore & Cao, 2008). Indeed, hemodynamics might modulate neural activity through direct and indirect mechanisms: for example, Moore & Cao propose that an upregulation of hemodynamic signals might increase the delivery of diffusible blood-borne messengers and modulate neural activity mechanically as well as thermically. Indirectly, changes in hemodynamics might regulate astrocytes, which themselves in turn would regulate neural activity. Support for such hypothesis is provided through studies using angiogenesis inhibitors demonstrating that angiogenesis but not neurogenesis is critical for learning (Kerr et al., 2010). The top five categories of differentially expressed genes experiencing a consistent downregulation across different learning stages, contained cardiovascular-related categories, translational regulation as well as changes in circadian rhythm. The latter seems to be of particular interest given previous reports of this category being affected during learning consolidation (Horowski et al., 2004; Smarr et al., 2014; Verwey et al., 2016) as well as an important factor in several neuropsychiatric disorders such as depression and schizophrenia (Walker et al., 2020). Furthermore, genes involved in circadian regulation have also been described in the context of the cognitive evolution, including human evolution (Fontenot et al., 2017).

The top five early-upregulated gene categories contained cellular and synaptic processes, which again were meaningful in the given context: inhibitory postsynaptic potential, regulation of glutamatergic synaptic transmission, protein refolding, regulation of cell migration and of developmental growth. These categories and certain genes of interest leading the group of differentially expressed genes within these categories, e.g. dopamine-2 receptors of A2A receptors (He et al., 2016) render the dataset meaningful in the context of cortico-basal ganglia related learning in health and disease. A first analysis of the top five gene categories with major down-regulation in particular during early learning stages further corroborate the molecular mapping of cortico-basal ganglia related learning in our assay. These categories again include learning and neurodevelopmental gene categories such as e.g. neural tube closure-related genes as well as molecular adaptations through e.g. mRNA slicing.

Taken together, data-driven analyses suggest a meaningful, major modulatory effect on learning-related and neuroplasticity-related gene clusters, categories and individual genes. The specific learning-related effect assessed through our experimental design renders our dataset useful for the exploitation in particular settings and contexts, including for example mouse models of neuropsychiatric or neurologic research with defects in cortico-basal ganglia-dependent learning and consolidation procedures.

References

Alexa, A., & Rahnenfuhrer, J. (2022). *topGO: Enrichment Analysis for Gene Ontology* (2.50.0) [Computer software]. Bioconductor version: Release (3.16). https://doi.org/10.18129/B9.bioc.topGO

Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience*, *9*, 357–381. https://doi.org/10.1146/annurev.ne.09.030186.002041

ANDREWS, S. (n.d.). *FastQC: a Quality Control Tool for High Throughput Sequence Data [Software]. 2010-2010. – ScienceOpen*. Retrieved December 30, 2022, from https://www.scienceopen.com/book?vid=a25d087f-b0ac-46c9-b031-6ecb0d873fd1

Balcombe, J. P., Barnard, N. D., & Sandusky, C. (2004). Laboratory routines cause animal stress. *Contemporary Topics in Laboratory Animal Science*, *43*(6), 42–51.

Benzina, N., N'Diaye, K., Pelissolo, A., Mallet, L., & Burguière, E. (2021). A cross-species assessment of behavioral flexibility in compulsive disorders. *Communications Biology*, *4*(1), 1–12. https://doi.org/10.1038/s42003-020-01611-y

Calkins, M. W. (1896). Association. An essay analytic and experimental. *The Psychological Review: Monograph Supplements*, *1*, i–56. https://doi.org/10.1037/h0092984

Cox, J., & Witten, I. B. (2019). Striatal circuits for reward learning and decision-making. *Nature Reviews. Neuroscience*, *20*(8), 482–494. https://doi.org/10.1038/s41583-019-0189-2

Dobin, A., & Gingeras, T. R. (2015). Mapping RNA-seq Reads with STAR. *Current Protocols in Bioinformatics*, *51*, 11.14.1-11.14.19. https://doi.org/10.1002/0471250953.bi1114s51

Dong, M., Thennavan, A., Urrutia, E., Li, Y., Perou, C. M., Zou, F., & Jiang, Y. (2021). SCDC: Bulk gene expression deconvolution by multiple single-cell RNA sequencing references. *Briefings in Bioinformatics*, *22*(1), 416–427. https://doi.org/10.1093/bib/bbz166

Fontenot, M. R., Berto, S., Liu, Y., Werthmann, G., Douglas, C., Usui, N., Gleason, K., Tamminga, C. A., Takahashi, J. S., & Konopka, G. (2017). Novel transcriptional networks regulated by CLOCK in human neurons. *Genes & Development*, *31*(21), 2121–2135. https://doi.org/10.1101/gad.305813.117

Gibson, E. M., Purger, D., Mount, C. W., Goldstein, A. K., Lin, G. L., Wood, L. S., Inema, I., Miller, S. E., Bieri, G., Zuchero, J. B., Barres, B. A., Woo, P. J., Vogel, H., & Monje, M. (2014). Neuronal activity promotes oligodendrogenesis and adaptive myelination in the mammalian brain. *Science (New York, N.Y.)*, *344*(6183), 1252304. https://doi.org/10.1126/science.1252304

Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. *Annual Review of Neuroscience*, *31*, 359–387. https://doi.org/10.1146/annurev.neuro.29.051605.112851

He, Y., Li, Y., Chen, M., Pu, Z., Zhang, F., Chen, L., Ruan, Y., Pan, X., He, C., Chen, X., Li, Z., & Chen, J.-F. (2016). Habit Formation after Random Interval Training Is Associated with Increased Adenosine A2A Receptor and Dopamine D2 Receptor Heterodimers in the Striatum. *Frontiers in Molecular Neuroscience*, *9*. https://www.frontiersin.org/articles/10.3389/fnmol.2016.00151

Hebb, D. O. (1949). *The Organization of Behavior: A Neuropsychological Theory*. Wiley.

Hintiryan, H., Foster, N. N., Bowman, I., Bay, M., Song, M. Y., Gou, L., Yamashita, S., Bienkowski, M. S., Zingg, B., Zhu, M., Yang, X. W., Shih, J. C., Toga, A. W., & Dong, H.-W. (2016). The mouse cortico-striatal projectome. *Nature Neuroscience*, *19*(8), 1100–1114. https://doi.org/10.1038/nn.4332

Ho, J., Tumkaya, T., Aryal, S., Choi, H., & Claridge-Chang, A. (2019). Moving beyond P values: Data analysis with estimation graphics. *Nature Methods*, *16*(7), 565–566. https://doi.org/10.1038/s41592-019-0470-3

Hoffman, G. E., & Roussos, P. (2021). Dream: Powerful differential expression analysis for repeated measures designs. *Bioinformatics (Oxford, England)*, *37*(2), 192–201. https://doi.org/10.1093/bioinformatics/btaa687

Hoffman, G. E., & Schadt, E. E. (2016). variancePartition: Interpreting drivers of variation in complex gene expression studies. *BMC Bioinformatics*, *17*(1), 483. https://doi.org/10.1186/s12859-016-1323-z

Horowski, R., Benes, H., & Fuxe, K. (2004). Striatal plasticity and motor learning-importance of circadian rhythms, sleep stages and dreaming.. *Parkinsonism & Related Disorders*, *10*(5), 315–317. https://doi.org/10.1016/j.parkreldis.2004.03.005

Janjic, A., Wange, L. E., Bagnoli, J. W., Geuder, J., Nguyen, P., Richter, D., Vieth, B., Vick, B., Jeremias, I., Ziegenhain, C., Hellmann, I., & Enard, W. (2022). Prime-seq, efficient and powerful bulk RNA sequencing. *Genome Biology*, *23*(1), 88. https://doi.org/10.1186/s13059-022-02660-8

Kaas, J. H., Merzenich, M. M., & Killackey, H. P. (1983). The reorganization of somatosensory cortex following peripheral nerve damage in adult and developing mammals. *Annual Review of Neuroscience*, *6*, 325–356. https://doi.org/10.1146/annurev.ne.06.030183.001545

Kassambara, A., & Mundt, F. (2020). *factoextra: Extract and Visualize the Results of Multivariate Data Analyses* (1.0.7) [Computer software]. https://CRAN.R-project.org/package=factoextra

Kerr, A. L., Steuer, E. L., Pochtarev, V., & Swain, R. A. (2010). Angiogenesis but not neurogenesis is critical for normal learning and memory acquisition. *Neuroscience*, *171*(1), 214–226. https://doi.org/10.1016/j.neuroscience.2010.08.008

Kolde, R. (2019). *pheatmap: Pretty Heatmaps* (1.0.12) [Computer software]. https://CRAN.R-project.org/package=pheatmap

Lamothe, H., Schreiweis, C., Lavielle, O., Mallet, L., & Burguiere, E. (2021). Not only compulsivity: The SAPAP3-KO mouse reconsidered as a comorbid model expressing a spectrum of pathological repetitive behaviors. *BioRxiv*, 2020.01.22.915215. https://doi.org/10.1101/2020.01.22.915215

Liao, Y., Smyth, G. K., & Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Research*, *47*(8), e47. https://doi.org/10.1093/nar/gkz114

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, *17*(1), 10–12. https://doi.org/10.14806/ej.17.1.200

McGeorge, A. J., & Faull, R. L. (1989). The organization of the projection from the cerebral cortex to the striatum in the rat. *Neuroscience*, *29*(3), 503–537.

Merzenich, M. M., Kaas, J. H., Wall, J., Nelson, R. J., Sur, M., & Felleman, D. (1983). Topographic reorganization of somatosensory cortical areas 3b and 1 in adult monkeys following restricted deafferentation. *Neuroscience*, *8*(1), 33–55. https://doi.org/10.1016/0306-4522(83)90024-6

Moore, C. I., & Cao, R. (2008). The hemo-neural hypothesis: On the role of blood flow in information processing. *Journal of Neurophysiology*, *99*(5), 2035–2047. https://doi.org/10.1152/jn.01366.2006

Pan, W. X., Mao, T., & Dudman, J. T. (2010). Inputs to the Dorsal Striatum of the Mouse Reflect the Parallel Circuit Architecture of the Forebrain. *Frontiers in Neuroanatomy*, *4*. https://doi.org/10.3389/fnana.2010.00147

Parekh, S., Ziegenhain, C., Vieth, B., Enard, W., & Hellmann, I. (2018). zUMIs—A fast and flexible pipeline to process RNA sequencing data with UMIs. *GigaScience*, *7*(6), giy059. https://doi.org/10.1093/gigascience/giy059

R Foundation for Statistical Computing, Vienna, Austria. (n.d.). *R Core Team (2022). R: a language and environment for statistical computing*. Retrieved December 30, 2022, from https://www.gbif.org/tool/81287/r-a-language-and-environment-for-statistical-computing

Robinson, M. D., & Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology*, *11*(3), R25. https://doi.org/10.1186/gb-2010-11-3-r25

Roy, N. A., Bak, J. H., Akrami, A., Brody, C. D., & Pillow, J. W. (2021). Extracting the dynamics of behavior in sensory decision-making experiments. *Neuron*, *109*(4), 597-610.e6. https://doi.org/10.1016/j.neuron.2020.12.004

Schaum, N., Karkanias, J., Neff, N. F., May, A. P., Quake, S. R., Wyss-Coray, T., Darmanis, S., Batson, J., Botvinnik, O., Chen, M. B., Chen, S., Green, F., Jones, R. C., Maynard, A., Penland, L., Pisco, A. O., Sit, R. V., Stanley, G. M., Webber, J. T., … Principal investigators. (2018). Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature*, *562*(7727), 367–372. https://doi.org/10.1038/s41586-018-0590-4

Skinner, B. F. (1963). Operant behavior. *American Psychologist*, *18*, 503–515. https://doi.org/10.1037/h0045185

Smarr, B. L., Jennings, K. J., Driscoll, J. R., & Kriegsfeld, L. J. (2014). A Time to Remember: The Role of Circadian Clocks in Learning and Memory. *Behavioral Neuroscience*, *128*(3), 283–303. https://doi.org/10.1037/a0035963

Thorn, C. A., & Graybiel, A. M. (2014). Differential entrainment and learning-related dynamics of spike and local field potential activity in the sensorimotor and associative striatum. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(8), 2845–2859. https://doi.org/10.1523/JNEUROSCI.1782-13.2014

Thorndike, E. L. (1927). The law of effect. *The American Journal of Psychology*, *39*, 212–222. https://doi.org/10.2307/1415413

Verwey, M., Dhir, S., & Amir, S. (2016). Circadian influences on dopamine circuits of the brain: Regulation of striatal rhythms of clock gene expression and implications for psychopathology and disease. *F1000Research*, *5*, F1000 Faculty Rev-2062. https://doi.org/10.12688/f1000research.9180.1

Walker, W. H., Walton, J. C., DeVries, A. C., & Nelson, R. J. (2020). Circadian rhythm disruption and mental health. *Translational Psychiatry*, *10*(1), 1–13. https://doi.org/10.1038/s41398-020-0694-0

White, N. M., & McDonald, R. J. (2002). Multiple parallel memory systems in the brain of the rat. *Neurobiology of Learning and Memory*, *77*(2), 125–184. https://doi.org/10.1006/nlme.2001.4008

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wilke, C. O. (2020). *cowplot: Streamlined Plot Theme and Plot Annotations for "ggplot2"* (1.1.1) [Computer software]. https://CRAN.R-project.org/package=cowplot

Yin, H. H., & Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nature Reviews. Neuroscience*, *7*(6), 464–476. https://doi.org/10.1038/nrn1919

Yin, H. H., Mulcare, S. P., Hilário, M. R. F., Clouse, E., Holloway, T., Davis, M. I., Hansson, A. C., Lovinger, D. M., & Costa, R. M. (2009). Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nature Neuroscience*, *12*(3), 333–341. https://doi.org/10.1038/nn.2261

Data Availability

Raw expression data will be submitted to Array Express (https://www.ebi.ac.uk/biostudies/arrayexpress).

Contributions

CS, WE, and EB conceptualised the research; EL conducted experiments; EL, ZK, IH, WE, CS, EB analysed data; EL, ZK, CS, WE wrote the article; CS, WE, EB edited the manuscript.

Acknowledgements

All animal work was conducted at the PHENO-ICMice facility. This work furthermore benefitted from the equipment and services from the iGenSeq core facility at ICM for the genotyping of the animals. This work was realised with the following funding: Agence Nationale de la Recherche (ANR-19-ICM-DOPALOOPS) (EB, WE), the Fondation pour la Recherche Medicale (FDT202204015143) (EL), and the L'Oréal-UNESCO Fellowship 2016 (CS). The core facilities were supported by “Investissements d’avenir” (ANR-10-IAIHU-06 and ANR-11-INBS-0011-NeurATRIS) and “Fondation pour la Recherche Médicale”.

Conflict of interest.

The authors declare no conflict of interest.

Figures with legends

Figure 1

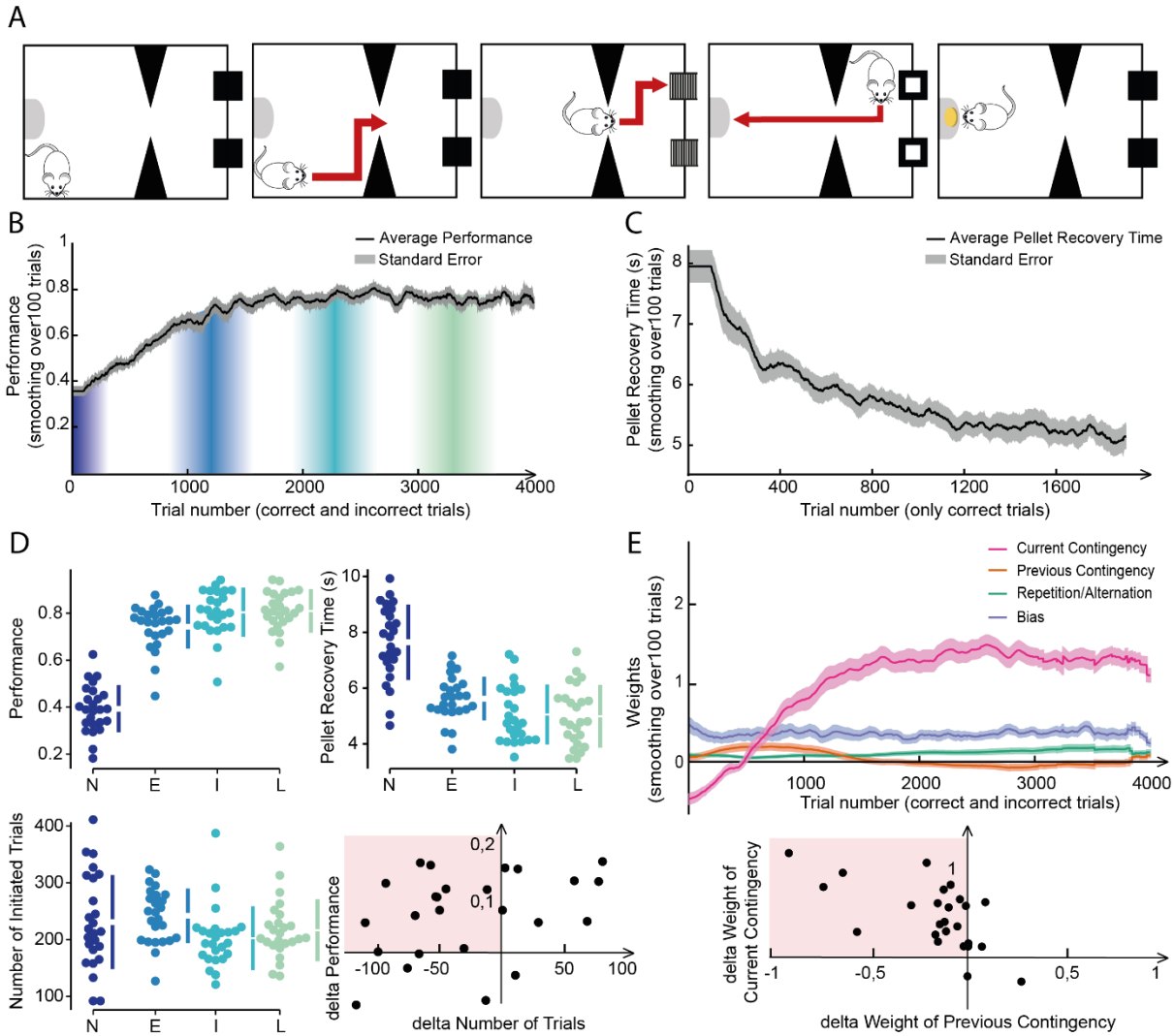

Figure 1. A self-motivated, visual discrimination task to capture learning and consolidation. A – Schematic illustration of the custom-adapted operant conditioning chambers (top-view). The animal lives inside the operant chamber, which is equipped with a food pellet dispenser at the rear end and two touchscreens displaying the task-relevant stimuli at the front end. The animal triggers trial onset by crossing the laser-beam equipped gate. B – Performance follows a sigmoid learning curve across up to 4000 trials. Plotted is the mean and SEM of n = 25 wildype animals. C – Pellet recovery times decreases across learning stages and serves as an additional behavioural readout across different learning stages. D – Individual performances, pellet recovery times and number of initiated trials during the night when reaching each learning stage. The bottom right graph represents the difference of late versus early learning stages in number of trials (x-axis) over the difference of performance (y-axis), and demonstrates gained efficiency across learning through the clustering of most animals in the top left quadrant. E –

Relative weights of PsyTrack several hidden strategies (Roy et al., 2021) that were available to the animals during the task. Pink: weight of applying current contingency, i.e., the animal follows stimulus-outcome contingencies in the current trial; orange: weight of applying previous contingency, i.e., the animal applies a win-stay/loose-shift strategy; green: general repetition/alternation strategy; violet: general lateralization bias. The bottom graph represents the differences between late and early learning stages in the weights of applying previous (x-axis) over current contingencies (y-axis), and shows a shift of the majority of the animals away from win-stay/loose-shift towards the optimal current contingency strategies. Plots in panels B-D represent means +/- SEM. The colour code in panels B and D represents the naïve, early, intermediate and late learning stages in dark blue, blue, turquoise, and green, respectively. N: naïve, E: early, I: intermediate, L: late learning stages.

Figure 2

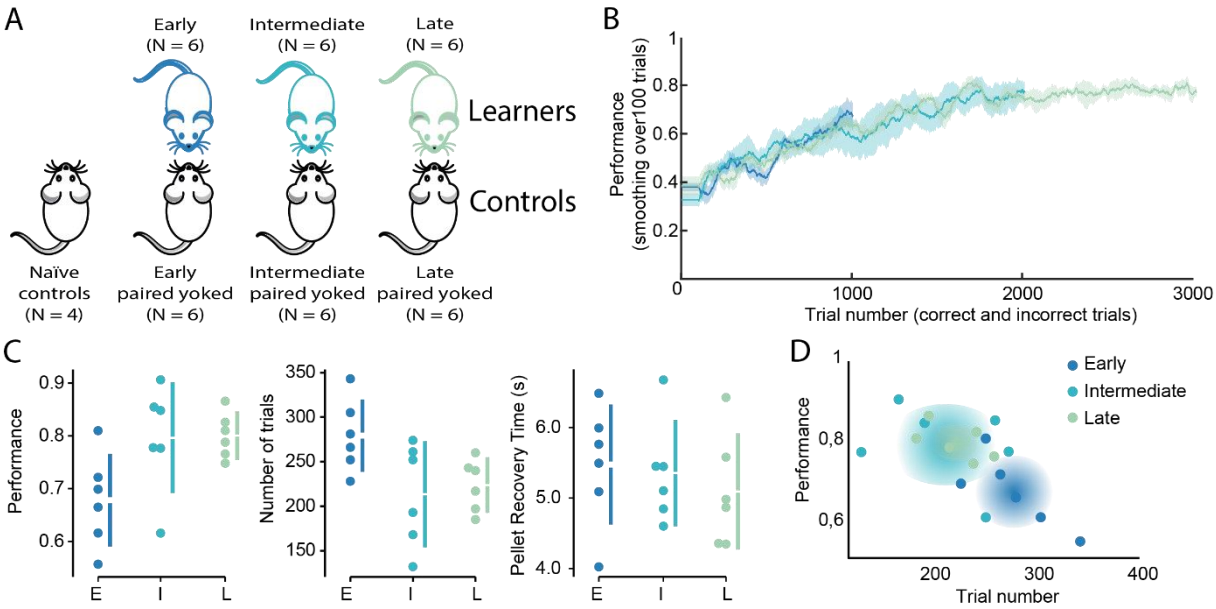

Figure 2. Replication of behavioural readout parameters in animals applied for molecular mapping. A – Illustration of experimental design. Groups mice of each learning stage (early, intermediate, late) were paired with yoked controls. Naïve mice were added as outgroup controls for general experimental procedures. B – Average performance (mean +/- SEM) of each learning stage group until stage criterion: 1000, 2000, and 3000 trials for early, intermediate, and late learners, respectively. C – Individual data points of performance, number of initiated trials, and pellet recovery times as behavioural readouts of the night prior to reaching stage criterion. D – Trial number (x-axis) over performance for individual animals of each learning stage during the night prior to reaching learning stage criterion. Early, intermediate and late learners are colour coded in blue turquoise, and green, respectively. E: early, I: intermediate, L: late.

Figure 3

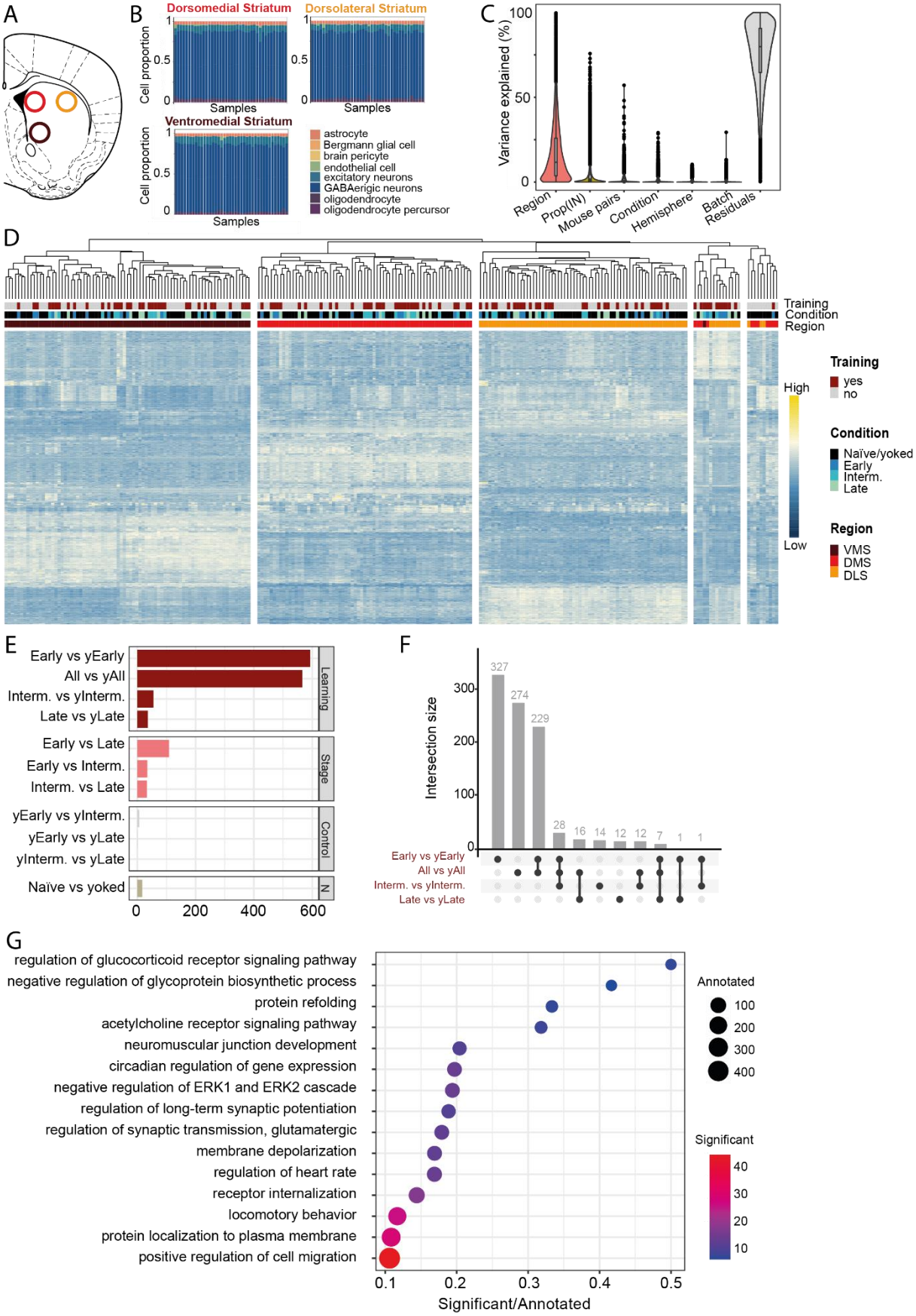
A
B
Dorsomedial Striatum
Dorsolateral Striatum
Ventromedial Striatum
Cell proportion
Samples
astrocyte
Bergmann glial cell
brain pericyte
endothelial cell
excitatory neurons
GABAerigic neurons
oligodendrocyte
oligodendrocyte percursor
C
Variance explained (%)
Region
Prop(IN)
Mouse pairs
Condition
Hemisphere
Batch
Residuals
D
Training
Condition
Region
High
Low
Training
yes
no
Condition
Naïve/yoked
Early
Interm.
Late
Region
VMS
DMS
DLS
E
Early vs yEarly
All vs yAll
Interm. vs yInterm.
Late vs yLate
Learning
Early vs Late
Early vs Interm.
Interm. vs Late
Stage
yEarly vs yInterm.
yEarly vs yLate
yInterm. vs yLate
Control
Naïve vs yoked
N
F
Intersection size
327
274
229
28
16
14
12
12
7
1
1
Early vs yEarly
All vs yAll
Interm. vs yInterm.
Late vs yLate
G
regulation of glucocorticoid receptor signaling pathway
negative regulation of glycoprotein biosynthetic process
protein refolding
acetylcholine receptor signaling pathway
neuromuscular junction development
circadian regulation of gene expression
negative regulation of ERK1 and ERK2 cascade
regulation of long-term synaptic potentiation
regulation of synaptic transmission, glutamatergic
membrane depolarization
regulation of heart rate
receptor internalization
locomotory behavior
protein localization to plasma membrane
positive regulation of cell migration
Significant/Annotated
Annotated
Significant

Figure 3. Learning-related differential gene expression is strongest during early learning.

A – Schematic illustration of striatal sample collection. Brown: limbic, ventromedial striatum, red: associative, dorsomedial striatum, yellow: sensorimotor, dorsolateral striatum. B – All striatal samples show expected cell type composition. C – Distribution of variance explained by experimental factors that were included in the GLMM analyses of differential gene expression. D – Heatmap of hierarchically clustered gene expression. Most differential gene expression could be explained by regional sampling; remaining gene expression patterns mostly cluster into learning versus non-learning condition. Yellow/blue colour spectrum represents up- or downregulated expression levels. Samples are colour-coded according to condition (learner versus yoked groups), learning stage (naïve, early, intermediate, late) and regional sampling (ventromedial, dorsomedial, dorsolateral striatum). E – Pooled differential expression analysis showed major learning-related changes especially during early stages. Yoked and naïve outgroup control groups with nearly absent differential gene expression demonstrate learning-specific expression in our task design. F – Comparison of expression changes across learning and individual learning stages. G – 15 most highly enriched TopGO gene ontology categories of pooled learning-related expression changes.

Figure 4

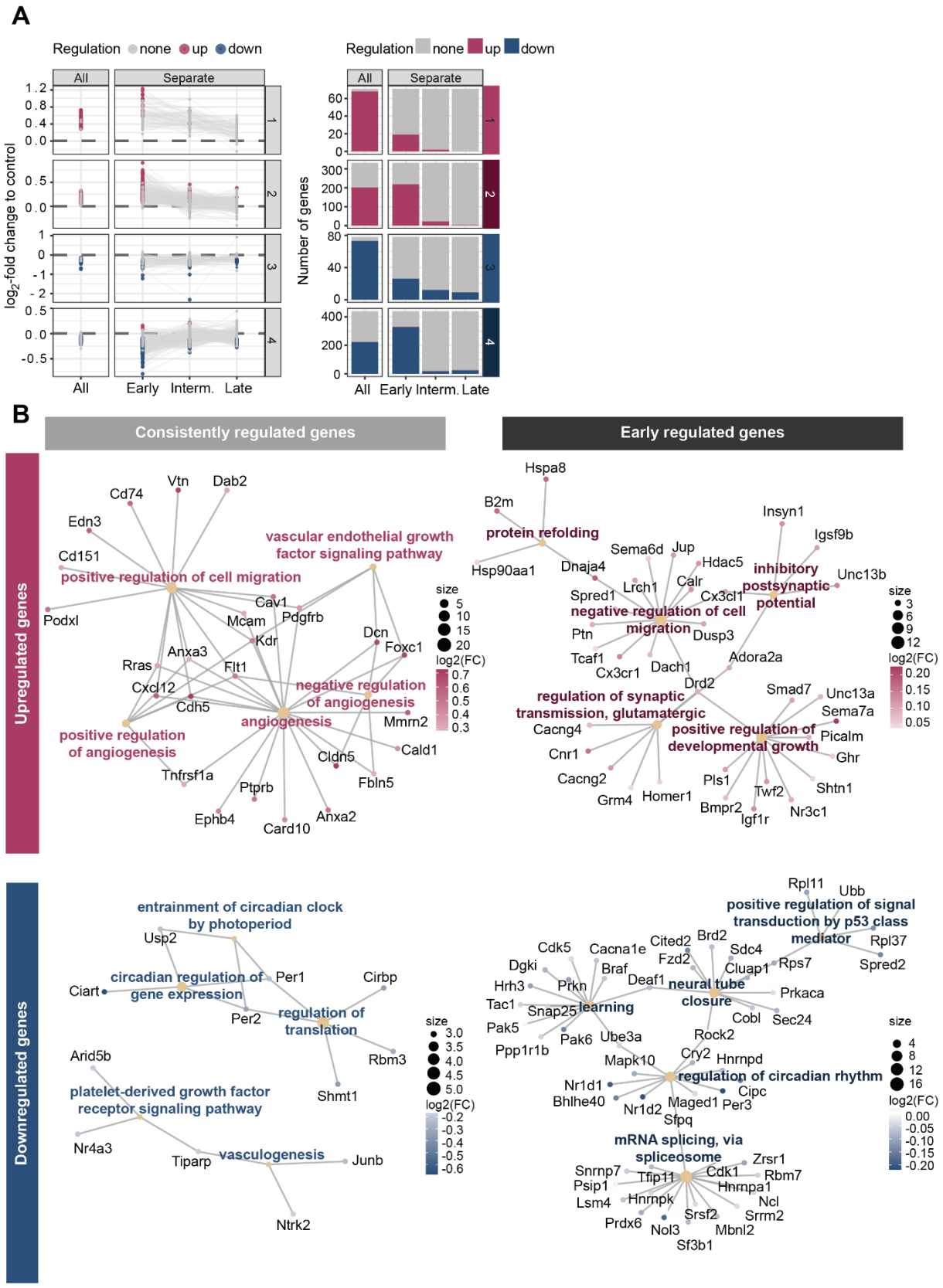
A
Regulation none up down
All
Separate
log2-fold change to control
Number of genes
All Early Interm. Late
B
Consistently regulated genes
Early regulated genes
Upregulated genes
Downregulated genes
vascular endothelial growth factor signaling pathway
positive regulation of cell migration
negative regulation of angiogenesis
angiogenesis
positive regulation of angiogenesis
protein refolding
inhibitory postsynaptic potential
negative regulation of cell migration
regulation of synaptic transmission, glutamatergic
positive regulation of developmental growth
entrainment of circadian clock by photoperiod
circadian regulation of gene expression
regulation of translation
platelet-derived growth factor receptor signaling pathway
vasculogenesis
positive regulation of signal transduction by p53 class mediator
neural tube closure
learning
regulation of circadian rhythm
mRNA splicing, via spliceosome
size
log2(FC)

Figure 4. Learning-regulated expression changes cluster into consistently versus early-learning enriched expression changes.

A – Learning-regulated differentially expressed genes cluster into four categories when plotted by their expression level across early, intermediate, and late learning stages: consistently up- or downregulated and mostly early-learning up- and downregulated genes. B – Cnet plots of the five most enriched gene ontology categories of each cluster with their respective differentially regulated genes, colored by their log2-fold change relative to the respective yoked control. Pink and blue colour coding represents differential up- or downregulated expression patterns. Light and dark grey represent consistently versus early-regulated differential gene expression.