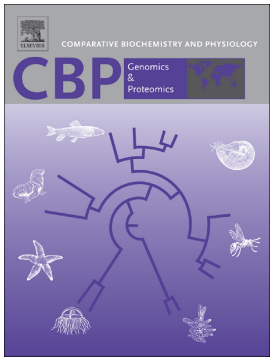
Journal Pre-proof



DNA methylation variation in the brain of laying hens in relation to differential behavioral patterns

Carlos Guerrero-Bosagna, Fábio Pértille, Yamenah Gomez, Shiva Rezaei, Sabine Gebhardt, Sabine Vögeli, Ariane Stratmann, Bernhard Vöelkl, Michael J. Toscano

PII: S1744-117X(20)30047-2

DOI: https://doi.org/10.1016/j.cbd.2020.100700

Reference: CBD 100700

To appear in: *Comparative Biochemistry and Physiology - Part D: Genomics and Proteomics*

Received date: 12 January 2020

Revised date: 27 April 2020

Accepted date: 17 May 2020

Please cite this article as: C. Guerrero-Bosagna, F. Pértille, Y. Gomez, et al., DNA methylation variation in the brain of laying hens in relation to differential behavioral patterns, *Comparative Biochemistry and Physiology - Part D: Genomics and Proteomics* (2018), https://doi.org/10.1016/j.cbd.2020.100700

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

source: https://doi.org/10.7892/boris.145123 | downloaded: 12.10.2024

© 2018 Published by Elsevier.

DNA methylation variation in the brain of laying hens in relation to differential behavioral patterns

Carlos Guerrero-Bosagna[1*], Fábio Pértille[1,2], Yamenah Gomez[3], Shiva Rezaei[1], Sabine Gebhardt[3], Sabine Vögeli[3], Ariane Stratmann[3], Bernhard Vöelkl[3], Michael J. Toscano[3]

1-Avian Behavioural Genomics and Physiology Group, IFM Biology, Linköping University, Linköping, Sweden, 58 183.

2-Animal Biotechnology Laboratory, Animal Science Department, University of São Paulo (USP)/ "Luiz de Queiroz" College of Agriculture (ESALQ), Piracicaba, São Paulo, Brazil.

3- Center for Proper Housing: Poultry and Rabbits (ZTHZ), Division of Animal Welfare, VPH Institute, University of Bern, Burgerweg 22, 3052, Zollikofen, Switzerland

*Corresponding author, email: carlos.guerrero.bosagna@liu.se

Journal Pre-proof

Abstract

Domesticated animals are unique to investigate the contribution of genetic and non-genetic factors to specific phenotypes. Among non-genetic factors involved in phenotype formation are epigenetic mechanisms. Here we aimed to identify whether relative DNA methylation differences in the nidopallium between groups of individuals are among the non-genetic factors involved in the emergence of differential behavioral patterns in hens. The nidopallium was selected due to its important role in complex cognitive function (i.e., decision making) in birds. Behavioral patterns that spontaneously emerge in hens living in a highly controlled environment were identified with a unique tracking system that recorded their transitions between pen zones. Behavioral activity patterns were characterized through three classification schemes: (i) daily specific features of behavioral routines (Entropy), (ii) daily spatio-temporal activity patterns (Dynamic Time Warping), and (iii) social leading behavior (Leading Index). Unique differentially methylated regions (DMRs) were identified between behavioural patterns emerging within classification schemes, with entropy having the higher number. Functionally, DTW had double the proportion of affected promoters and half of the distal intergenic regions. Pathway enrichment analysis of DMR-associated genes revealed that Entropy relates mainly to cell cycle checkpoints, Leading Index to mitochondrial function, and DTW to gene expression regulation. Our study suggests that different biological functions within neurons (particularly in the nidopallium) could be responsible for the emergence of distinct behaviour patterns and that epigenetic variation within brain tissues would be an important factor to explain behavioral variation.

Keywords: Chickens, Behavior, Epigenetics, Brain, DNA methylation

Journal Pre-proof

Introduction

Measuring the quantitative contribution of genetic composition is usually the first step to understand the relative importance genes play in behavior [1]. Towards this goal, domesticated animals provide unique models to investigate the contribution of genetic and non-genetic factors to specific phenotypes. Commercial poultry offer a particularly relevant model to gain this understanding as there is normally no maternal role once the egg is laid, excluding a major confounding factor, as maternal care is shown to influence DNA methylation in the offspring in mammals [2-4]. Variation beyond that attributed to genetic factors can be in part explained by epigenetic mechanisms and the role they exert in the formation of phenotypes. Epigenetic mechanisms involve chemical modifications of the DNA that regulate gene expression and can be maintained after cell divisions [5]. Epigenetic mechanisms are, on one hand sensitive to environmental influences, and, on the other hand, fundamental players in shaping the adult phenotype of individuals [6].

From a neurobiological perspective, epigenetic mechanisms are reported to be involved in processes such as memory, cognition, synaptic plasticity [7] and regulation of stress response [8].

Regarding efforts to understand origins and regulation of behavioral variation, the concept of individuality has become an important research topic across many species [9, 10], including poultry. Within commercial systems, chickens were recently shown to manifest highly consistent movement and location patterns [11]. Despite the growing interest in this theme, the causal biological mechanisms that lead to behavioral variation are poorly understood, particularly in large groups typical of commercial livestock such as laying hens. At a fundamental level, the emergence of individual behavior phenotypes within a population will involve both genetic and non-genetic mechanisms as with other phenotypes. In domesticated animals, behavioral traits are associated with low to moderate heritability [12-14]. In modern commercial laying hen hybrids, the component of the heritability attributed to social traits is estimated to be at the same level of direct effects, i.e., 30% [15].

Epigenetic variation is one of the biological mechanisms that explain non-genetic heritability [16, 17] and is thought to account for a sizeable part of phenotypic variability [18]. The best studied epigenetic mechanism is DNA methylation, which involves the enzymatic addition of methyl groups (-CH_3) to 5' to 3' oriented CG dinucleotides, known as CpG sites [19]. In humans, differential methylation in neurons explain great part of the heritability of neurological disorders such as schizophrenia, addictive behavior, and neuroticism [20]. In the present study we aimed to identify whether DNA methylation changes in the brain (nidopallium) are among the non-genetic factors involved in the emergence of differential behavioral patterns in chickens. These behavioral patterns were identified with a unique tracking system that involves recording transitions between specific zones of a pen. The nidopallium was selected as the brain area for epigenetic analysis due to its important role in complex cognitive function in the avian brain [21]. This brain region was selected because it controls decision-making, which is a relevant neurological process that takes place when the animals weight the attributes of each zone in the pen (e.g., the relative amount of natural versus artificial lighting in outside or inside areas.).

We characterized behavioral traits and activity levels using three types of classification schemes that our group has been employing in a variety of contexts [11, 22]. These schemes were: (i) daily specific features of behavioral routines and their associated stability (Entropy), (ii) daily spatio-temporal activity patterns (Dynamic Time Warping), and (iii) a study-wide assessment of social leading behavior (Leading Index). In the present study we investigated the relationship between DNA methylation in the brain of chickens (nidopallium) and behavioral patterns detected through these methods. We identified a number of Differentially Methylated Regions (DMRs) associated with distinct patterns of behavioral activity that spontaneously emerged in our chicken population. Furthermore, we explored how the identified DNA methylation changes could influence genomic expression in each case.

Journal Pre-proof

Materials and Methods

Animals and housing

Beginning at one day of age until 17 weeks of age, 2,840 commercial Brown Nick laying hens (https://www.hn-int.com/eng/commercial-layers/brownnick.php, accessed 12-9-2019) were housed at the Aviforum (Zollikofen, CH), a contract research facility focusing on commercial poultry. Animals were kept in eight pens (355 hens/pen) of a rearing barn equipped with one of two aviary systems (four pens with Inauen Natura, Inauen AG, Appenzell, Switzerland, and four pens with Landmeco Harmony, Globogal AG, Lenzburg, Switzerland). Each pen had floors that were covered with wood shavings. The aviary system contained round metal perches, automatic feeders, nipple drinkers, and manure belts. The chicks had access to a covered outdoor area (winter-garden) from six weeks of age onwards.

At 17 weeks of age, birds were transferred to an on-site commercial laying hen house that was divided into two halves of which only one side was used for the current study. Each pen was equipped with a system that allowed tracking of individual animals, described in more detail below. The four pens (12.9 m^2) contained a Rihs Bolegg II commercial aviary system (Krieger AG, Ruswil, Switzerland) with a stocking density of 9.33 hens/m^2. Birds from each rearing pen were distributed across each laying pen in a stratified manner. The barn interior included an aviary structure and group nests along one wall, with the floor covered with 10 cm of wood shavings. The aviary was 2.40 m high and consisted of three tiers with the following equipment integrated into the structure: manure belt, feeding chain, and nipple drinkers within the lowest tier; a manure belt within the middle tier; a feeding chain and nipple drinkers within the highest tier. Plastic mushroom-shaped perches were provided on the lowest and highest tiers. Plastic platforms to move between tiers were provided along both aviary sides (30 cm in width and at 70 cm height from the floor) and nest entries were square plastic grids (size 2.5 × 5 cm). External to the internal barn area, birds had access to a winter-garden (average size: 17.55 m^2; equipped with litter, nipple drinkers, perches), a fenced area

containing small stones (stone yard, average size per pen: 88 m^2), and a pasture ("free-range area", average size per pen: 288 m^2). The winter-garden consisted of an area entirely covered by a solid roof and surrounded by wire mesh on the sides and in between pens, thereby preventing birds from exiting the area with the exception of a pophole that could manually be opened. Each pen – including both external and internal areas - was separated by fencing to maintain divided populations. Within a pen, each area was further divided with fencing (or the barn wall) so access could be limited to the interior or outdoor areas as required by management protocol. Movement between areas (i.e. inside, winter-garden, stone yard, free-range) was via a single location between areas (pop hole or gate) that provided unobstructed access when opened. Transitions between areas could only occur between two juxtaposed areas, e.g., transitioning directly from the barn to the free-range areas without passing through the winter-garden and stone yard was not possible. Artificial light was provided in the barn from 0200 to 1700 h with a transitional phase of five min beginning at 0200 h and 15 min at 1645 h. Natural daylight was provided from 0800 to 1630 h through windows controlled by curtains or on pasture. Birds were encouraged into the barn interior around 16:30.

Movement and location acquisition

To record hen movement, a system similar to that described by Gebhardt-Henrich et al. [23] was used with some minor modifications. Within each pen 120 out of 355 hens (33.8%) were randomly selected to be fitted with Radio Frequency Identification tags (RFID, Hitag S 2048 bits, low frequency of 125 kHz, diameter: 4.0 mm, length: 34.0 mm) attached to leg bands. This random assignment of RFID tags to 120 hens for tracking was done on the day the barn was populated. Two sets of antennas (*Gantner Pigeon System* (http://www.benzing.cc/, accessed 10.06.19) were positioned immediately on either side of the transition points connecting two areas (e.g., barn/winter-garden) in a manner that hens transitioning had to pass over each set. In order to protect the antennas from weather and staff working in the area, they were entirely encased within a small wooden box that ran the length and both sides of each transition area.

The RFID system operated by registering and recording the date and time and date that individual RFID tags (worn by the hens) came within a vertical distance of 15 cm to an antenna. The inclusion of antennas on either side of the transition areas represents an added level of assurance as movement between two areas required registration of two events – both entrance into the area (e.g. registration inside the house followed by a transition into a second area, e.g. , the winter-garden). Collected data, including: unique tag identification number, timestamp (with a precision of 0.1 s), and antenna number, were written to a connected computer. The system allowed for multiple tags (and the associated hen) to be registered by the same antenna at the same time. The direction of movement was deduced from the order in which the antennas detected the tags. More details of the RFID system and its reliability are provided by Gebhardt-Henrich et al. [24].

Data was not analyzed for all days of the study period because access to the stone yard or free-range areas was restricted during poor weather or pasture maintenance. Tracking data were recorded only on 72 days (within the contiguous 166-day study period) when all pens were given access to all areas

Based on time-sampled transition data we calculated the following measures for each hen: time inside, time outside, number of transitions, entropy, order leaving, order returning, number of days outside, a leading index, as well as a classification of movement type based on Dynamic Time Warping (DTW). For these measures, the stone yard and free-range areas are considered outside. Time inside the barn is the proportion of time that the birds spend inside the barn between the opening and closing of the popholes/gates during all days. Time outside is the proportion of time that birds spend either in the stone yard or in the free-range area between the opening and closing of the popholes/gates during all days. Number of transitions is the total number of recorded transitions for a specific hen between opening and closing of the pophole/gates summed over all days. The entropy for each hen was calculated as the sample entropy, which is the negative natural logarithm of the conditional probability that two sequences similar for a number of points remain similar at the next

point, excluding self-matches [25] (Supplementary Figure S1). The order for leaving is based on the order in which hens of a pen entered the stone yard for the first time on a given day. Ordinal numbers for each hen are averaged over all observational days. The order for returning is based on the order in which hens of a pen returned from the stone yard and re-entered the winter garden for the last time on a given day. The numbers of days outside is the number of days where a bird was at least once recorded entering the stone yard or the free-range area. The leading index gives for each hen the proportion of all its transition where it was a leader, defined as being followed more closely in time by another bird than following another bird. For the differential methylation analysis both entropy and leading index were converted into binary measures by splitting the subsample of tissue sampled animals into two evenly sized groups with the median entropy and leading index values of the subset as the threshold values. DTW is a shape-based time series analysis, comparing the dissimilarity of two time series independent of their individual length [26, 27]. DTW creates a dissimilarity distance matrix including each pairwise comparison of time series and was calculated using the R-package dtwclust [28]. High DTW score exist when two time series are very dissimilar from each other. If pairwise comparisons of time series from a single bird have high DTW values, this indicates that the bird was not very consistent in its movement patterns. If DTW distance scores between birds are high, this indicates that the birds had rather different movement patterns. Hierarchical clustering analyses of these dissimilarity matrices was performed with the DIANA method (divise analysis, [19]). Beyond these behavioral classifications, no other categories were used to group or select animals.

Tissue collection and DNA extraction

The 24 animals using in this study were haphazardly selected from a larger population devoted for multiple parallel investigations in hens. Animals were sacrificed via intravenous injection with pentobarbital (Esconarkon, 0.3 ml/hen). Immediately thereafter, brains were removed from the skulls and transferred to petri dishes containing 0.1M PBS. The hemispheres were divided along the longitudinal fissure with a scalpel. The brain section used for the genomic and epigenomic analyses

was removed using scissors from the exposed area underneath the lateral ventricle and adjacent to the midline. As this region came from the caudal half of the forebrain it consisted primarily of caudal medial nidopallium [29]. The samples were then stored at -80°C until shipped on dry ice to Linköping University for further processing. After arriving, DNA was extracted from the brain sections using the DNeasy Blood and Tissue Kit from Qiagen, following the manufacturer's instructions.

Preparation of sequencing libraries

To prepare the sequencing library we used an approach that combines Genotyping by Sequencing (GBS) [30] and the Methylated DNA immunoprecipitation (MeDIP) [31] techniques. We recently employed this methodological combination in previous studies [32]. The method consists in first digesting the genome with the *PstI* restriction enzyme (Thermo Scientific) in a suitable range for Illumina sequencing [30]. Illumina sequencing barcodes are then ligated at each end of the digested DNA fragments, allowing the pool of DNA samples to be immunoprecipitated together. Each pooled DNA sample contained different barcodes identifying each individual. The methylated fraction of the DNA is captured by an anti-methyl-cytosine antibody (Diagenode) [31]. After this step, the methylated DNA is amplified using PCR, which is followed by clean-up of primer dimers and unbound adapters [33, 34] before the samples are sent for paired-end sequencing on the Illumina HiSeq2500 platform with a read length of 125 bp at the SNP&SEQ facilities of the SciLifeLab (Stockholm, Sweden).

Bioinformatic analyses

The CASAVA (Illumina) program was used for the initial processing of the samples by converting the ".bcl" (base call files) to ".fastq" extensions, which is compatible with programs used for reads alignment. The quality of the reads was checked using FastQC v.0.11.3 [35]. Quality trimming was performed in short read sequences during the data processing. For both SNP calling and methylation analyses, quality-trimmed reads were aligned against the chicken reference genome (*Gallus_gallus* 5.0, NCBI) with the Bowtie2 tool v.2-2.3.4.2 [36] using default parameters. The

coverage depth of each sequenced file was determined using SAMtools v.0.1.19 [37] with the "depth" option. For the identification of differential methylation regions (DMR), uncalled and low quality score bases were eliminated using the process_radtags function in the Stacks v.1.39 program [38]. Following the alignment, sequencing data for each individual were then assigned to one of the experimental groups. For the identification of significant DMRs, the animals were divided into three classification schemes: Entropy (0 and 1), Leading Index (0 and 1) and DTW clustering (1, 2 and 3). DMRs were calculated by performing pairwise comparisons between DNA methylation levels (observed in the reduced genomes of RBCs) from animals belonging to two of the categories defined within each classification scheme. The MEDIPS package from R [39] was used for basic data processing, quality controls, normalization, and identification of differential coverage regions using default parameters. The BSgenome.Ggallus.UCSC.galGal5 package from Bioconductor was used as the chicken reference genome within R environment. Quality control was carried out to confirm enrichment of the methylated fraction of the genome. This was performed by calculating the average enrichment score. Enrichment scores should be> 1, with values around 2 signalling very good enrichment for methylated DNA. The main idea of this approach is to verify the extent of CpG enrichment in the regions obtained compared to the reference genome. For this, the function counts the number of Cs, the number of Gs, the number of CpGs and the total number of bases within the reference genome. Subsequently, the function calculates the relative frequency of CpGs (relH) and the observed / expected ratio of CpGs in the reference genome (GoGe). In addition, the function performs the same calculation for DNA sequences underlying regions of interest. The final enrichment values result from dividing the relative frequency of CpGs (or the observed / expected value) in the regions of interest by the relative frequency of CpGs (or the observed / expected value) of the reference genome.

We used the same specific parameters for the MEDIP package as we previously reported [32]. However, the parameter of P=0.01 was used as the threshold for the detection of stacked reads to call DMRs. To define the ROI (regions of interest) to be analysed, we used a bed.file obtained from

the Model-based Analysis of ChIP-Seq data (MACS) peak calling program (https://github.com/taoliu/MACS/) [40], using default parameters. Macs2 allows that large methylated regions are not arbitrarily divided into smaller windows. Therefore, the analysis is "peak specific". MACS2 improves the spatial resolution of the predicted sites, uses a dynamic parameter to capture local biases in the genome and improves the robustness and specificity of the prediction, being strongly indicated for fold-enrichment experiments [41]. Three thresholds defined the genomic windows obtained by the DMR analyses: $P < 0.0005$ for describing genes related to significant DMRs, $P < 0.005$ for exploratory analysis of DRM-gene related, and $p<0.05$ for enrichment pathways. To identify DMRs ($P < 0.05$) overlapping across the classification schemes, we used the GenomicRanges package in R environment.

The significant DMRs obtained were annotated against the chicken reference genome (BSgenome.Ggallus.UCSC.galGal5) using annotatePeak function from the ChIPseeker package [42] in R environment. In this function, we used the gg_txdb (as the transcript metadata) from GenomicFeatures package and org.Gg.eg.db package as the annotation database for the chicken genome. For the identification of affected molecular function, cellular components and biological processes, we used the DMR-associated genes for a Gene Ontology (http://geneontology.org) analysis performed through the enrichGO function within the ChIPseeker package [42]. For the identification of enriched molecular interaction and reaction networks we used the Kyoto Encyclopaedia of Genes and Genomes (KEGG; https://www.genome.jp/kegg), which was run with the enrichKegg function also within the ChIPseeker package [42]. Also, selected DMR-associated genes (described in the results) were used as input in the web-based Genemania tool (https://genemania.org, using default parameters) to obtain extended gene networks [43].

Journal Pre-proof

Results

Basic movement

The tracking system successfully registered 1,219,658 transitions of 421 tracked hens (range:0-6339) across the included days of the study period. Due to missed data we could not determine the birds' positions at all times, though on average (mean), the location of a bird was known to us in 84.3% of the time. Utilization of the four available areas differed substantially. While ten birds (2.4%) never left the indoor area, other birds spent up to 87% of available time outside of the barn during which access was available (Supplementary Figure S2) and 33 birds (7.8%) never entered the free-range area.

The individual-level variables (time inside the barn, time outside the barn, number of transitions, entropy, order leaving, order returning, number of days outside and leading index) were partly correlated (Supplementary Table S1). A principal component analysis of the normalized z-scores (Supplementary Figure S3; Supplementary Table S2) gives a rotation where the first principal component explains 55% and the second component explains 13% of the overall variation. The first component has high loadings for time inside and order going outside, and low loadings for the number of transitions, entropy and the number of days seen outside. The second component has by far the strongest loading for the leading index. Entropy was highly correlated with the number of transitions per day (r = 0.95, CI95 = 0.94—0.96). The time animals spent in the barn was correlated with: the order first entering the stone yard (as part of the outdoor area) (r = 0.59, CI95 = 0.52—0.65; a high order number means that the bird left later than most other birds in the day) and negatively correlated with: entropy (r 0 -0.69, CI95 = -0.74— -0.64), number of transitions (r = -0.60, CI95 = -0.66— -0.53), order of going back inside in the afternoon (r = -0.53, CI95 = -0.60— -0.45), and number of days the birds were recorded outside (r= -0.64, CI95 = -0.69— -0.58). Thus, birds that spent less time in the barn were generally more active in terms of movements between areas. The Leading index was not strongly correlated with any of the variables describing the movement

patterns (e.g, number of transitions: r = 0.205, CI95 = 0.11—0.30; time inside the barn: r = -0.24, CI95 = -0.33— -0.15; order going outside the barn r = -0.26, CI95 = -0.35— -0.17. The median leading index was 0.496 (interquartile range: 0.463 - 0.530) and the median entropy was 0.018 (interquartile range: 0.012 - 0.024). DTW analysis allowed the creation of dissimilarity matrices by computing summed dissimilarity distances for all pairs of the 24 hens (across both pen 1 and pen 2). Hierarchical clustering analyses of these dissimilarity matrices, performed with the DIANA method (divise analysis [28]), suggests the existence of clusters of distinct movement patterns. Based on this clustering analysis, we were able to identify three clusters representing markedly differing movement patterns (henceforth DTW clusters; Figure 1). The categorization of each hen within each behavioral classification scheme is shown in Supplementary Table S3

Basic Sequencing Features

We sequenced a reduced and methylation enriched genomic fraction of 24 individuals across experimental groups. The average sequencing yield across individuals was 233.7±4.1 million of bps. The average sequencing depth was 56.3X±9.0X covering ~4.06 million unique bps, which corresponds to 0.4% of the whole chicken genome. Details about individual sequencing coverage per DNA sample can be found in the Supplementary Table S4.

In order to verify the enrichment for CpGs in our sequenced reads, we calculated a CpG 'enrichment score', which compares CpGs in the genomic regions covered by the sequenced reads against CpGs in the whole reference genome. An 'enrichment score' of 2.80 ± 0.22 was obtained based on 94,331 CpGs identified across individuals on average. This corresponds to approximately 1% of all CpGs within the whole *Gallus gallus* genome.

Differential Methylation Analysis Between Experimental Groups

DMRs (P≤0.05) were identified between i) animals with low and high entropy, ii) animals characterized as leaders or followers (according to the calculated 'Leading Index'), and iii) animals differentially classified into three DTW clusters. Between animals with high and low entropy 110

DMRs were identified (Figure 2A). Additionally, 25 DMRs were found between animals characterized as followers or leaders (Figure 2B). In relation to the DTW clusters, 15 DMRs were identified between DTW clusters 1 and 2 (Figure 2C), 6 DMRs were identified between DTW clusters 2 and 3 (Figure 2D), and 7 DMRs were identified between DTW clusters 1 and 3 (Figure 2E). A selection of the DMRs with the lowest P values (P≤0.005) is shown in Table 1 with their respective annotations. The number of DMRs unique to, or overlapping between, classification schemes is shown in Figure 3A. The unique and overlapping DMRs observed between DTW clusters is shown in Figure 3B.

	P value	Classification scheme	Genomic location	Width (bp)	CpGs	Annotation	Gene Id	Entrez ID	Gene Symbol	Gene Name	Treatment with hypermethylation	Possible Expression effect
1	0.00012	DTW cluster 2 vs 1	chr26:870542-871014	472	1	Intron	ENSGALG00000031122.1	424336	NTNG1	Netrin G1	DTW cluster 2	Increased in DTW cluster 2
2	0.00036	Entropy	chr4:10886922-10887074	152	1	Promoter	ENSGALG00000007211.5	419305	CDH22	Cadherin 22	Low entropy	Decreased in low Entropy
3	0.00075	Entropy	chr20:10877159-10877501	342	14	Intron	ENSGALG00000002374.5	424772	SLC9A9	Solute carrier family 9 member A9	Low entropy	Increase in low Entropy
4	0.00091	Entropy	chr2:82081027-82081519	492	11	Distal Intergenic	ENSGALG00000040186.2	NA	NA	NA	Low entropy	Decreased in low Entropy
5	0.00133	Entropy	chr19:7725868-7726161	293	21	Distal Intergenic	ENSGALG00000002892.5	423653	SLC22A15L	Solute carrier family 22 member 15-like	Low entropy	Decreased in low Entropy
6	0.00210	Entropy	chr15:7902061-7902648	587	27	Distal Intergenic	ENSGALG00000005415.5	403089	TEAD1	TEA domain transcription factor 1	Low entropy	Decreased in low Entropy
7	0.00210	Entropy	chr20:10489420-10489893	473	17	Exon	ENSGALG00000002663.5	424777	ATR	ATR serine/threonine kinase	Low entropy	Increased in low Entropy
8	0.00219	DTW cluster 3 vs 2	chr12:2458270-2458563	293	17	Promoter	ENSGALG00000002722.5	396135	MST1	Macrophage stimulating 1 (hepatocyte growth factor-like)	DTW cluster 3	Decreased in DTW cluster 3
9	0.00317	Entropy	chrZ:73604888-73605174	286	10	Intron	ENSGALG00000029598.2	NA	NA	NA	Low entropy	Increased in low Entropy
10	0.00338	Entropy	chr12:7918429-7918698	269	23	Distal Intergenic	ENSGALG00000041167.2	NA	NA	NA	Low entropy	Decreased in low Entropy
11	0.00353	DTW cluster 3 vs 2	chr24:2807083-2807962	879	12	Promoter	ENSGALG00000043955.1	NA	NA	NA	DTW cluster 3	Decreased in DTW cluster 3
12	0.00383	DTW cluster 3 vs 2	chr10:20047539-20048245	706	39	Promoter	ENSGALG00000036163.1	100858113	ZNF710	Zinc finger protein 710	DTW cluster 3	Decreased in DTW cluster 3

***Table 1**- Genomic features and annotation of the DMRs with the highest p values.*

An annotated list of all the DMRs identified (P<0.05) in the current study is available in Supplementary Table S5. Within each classification scheme, functional annotation of all DMRs was performed to identify the genomic functional regions where the DMRs found locate. DMRs in each classification scheme were associated with different patterns of functional genomic features (Figure 4). In total, there were 81 genes associated to the DMRs. Interestingly, the patterns of functional annotation differ substantially across the classification schemes.

Next, KEGG (enrichKegg) and Gene Ontology (enrichGO; biological processes) pathway analyses were performed using the 81 genes associated to the DMRs as input. However, we selected the subset of these genes associated with each classification scheme for the pathway analyses related to them. Overall, for the Kegg analysis 26 genes mapped to at least one pathway in the Kegg database, and for the GO analysis 9 genes mapped to at least one pathway in the OrgDb database. With both analyses we identified genes significantly enriched in pathways. The enrichKegg analysis revealed genes enriched in pathways in relation to each of the comparisons performed within classification schemes

(Padj≤0.1). Some of these are well studied genes, relevant for a variety of biological pathways, such as RAF1 (associated to DMRs found between DTW clusters 2 and 1), GSL (associated to DMRs found between DTW clusters 3 and 1) and TGFβ2 (associated to DMRs found between followers and leaders). Also of interest is the fact that three genes associated to DMRs found between low and high entropy (RFWD2, ATR and SHISA5) belonged to the p53 signaling pathway (Table 2). The complete output of the enrichKegg analysis is shown in Supplementary Table S6. The enrichGO analysis revealed DMR-associate genes (Padj≤0.1) emerging within the Leading index classification scheme (RAF1), and when comparing DTW 2 vs 1 (TGF β2). These genes were also found in the enrichKegg analysis and are important for a number of biological process, as previously mentioned and as shown in the Supplementary Table S7.

Comparisson	DMR-Associated Gene enriched in Pathways	Pathways involved
DTW cluster 2 vs 1	ABCB6	ABC transporters
	FANCE	Fanconi anemia
	RAF1	VEGF signaling; VEGF signalling; ErbB signalling; Progesterone-mediated oocyte maturation; Gap junction; GnRH signaling; C-type lectin receptor signaling; Melanogenesis; Vascular smooth muscle contraction; Apelin signaling; FoxO signaling; Autophagy - animal; Apoptosis; Influenza A; mTOR signaling; Cellular senescence; Focal adhesion; Regulation of actin cytoskeleton
DTW cluster 3 vs 2	NTNG1	Cell adhesion molecules (CAMs)
	ATP2B4	Adrenergic signaling in cardiomyocytes; Calcium signaling
DTW cluster 3 vs 1	GSL	Arginine biosynthesis; Nitrogen metabolism; Glyoxylate and dicarboxylate metabolism; Alanine, aspartate and glutamate metabolism; Biosynthesis of amino acids; Necroptosis
Entropy: Low vs High	RFWD2	p53 signaling
	ATR	
	SHISA5	
Leading index: Follower vs Leader	TGFB2	TGF-beta signaling; AGE-RAGE signaling pathway in diabetic complications; Cell cycle; FoxO signaling; Cellular senescence
	AIFM1	Apoptosis; Necrosis
	ATP6V1AL	Phagosome; mTOR signaling

Table 2- *Gene enrichment obtained by KEGG pathways based on genes associated to significant DMR (P<0.05) found among behavioral patterns identified within each classification scheme.*

After this analysis, for each classification scheme we selected i) all the genes associated to the most highly significant DMRs (Table 1), and ii) those genes that were enriched in the Kegg pathway analysis. This subset of highly relevant genes was used to build extended gene networks and identify their biological impact using the web based GeneMania (https://genemania.org) tool [43]. The extended gene networks related to each classification scheme are shown in Figure 5A. The functional biological pathways significantly impacted by these extended gene network modules are shown in Supplementary Table S6. In order to better visualize the different biological functions affected by the extended gene networks associated to DMRs within each classification scheme we built word clouds using the terms of the abovementioned functional pathways (Figure 5B).

Discussion

Overview

In the current work we used a relatively large and uniform population of laying hens with the aim of identifying the relationship between behavioral patterns (detected by applying metrics in novel ways) and epigenetic variation (DNA methylation) in a brain region involved in the processing of higher cognitive abilities (nidopallium) in birds. Behavioral patterns that spontaneously emerged were able to be documented in our population of hens and were linked to specific DMRs. Genes associated to these DMRs are relevant from a neurobiological perspective but are also involved in other biological functions, notably cell cycle checkpoints and exit to repair.

This study is unique in several ways that contrast with previous efforts investigating interactions between behavior and the epigenome. Firstly, the study population of commercial laying hens is highly homogeneous in terms of both environmental exposures and genetic composition. All hens were hatched, reared, and housed together where they would be exposed to the same nutrition, lighting, climate, and other environmental factors, and without any possibility for maternal interactions. In terms of genetic variation, these farm animals are considered fairly homogenous due to intensive commercial breeding efforts to yield highly productive and feed efficient animals.

A second unique aspect of the current work is that behavioral data was collected longitudinally over an extended period of time within the animals' home pen which minimized disturbances to the animal. We believe the methods to track and analyze behavior used in the current study ensure the identified relationships are more indicative of actual commercial populations relative to efforts that utilize experimentally induced behaviors within specific testing paradigms conducted outside the home area. Although such approaches can serve as useful proxies for underlying behavioral traits [44], they present limitations when investigating longitudinal changes [45].

Another unique aspect is that data was collected in a relatively large group of animals where the utilization of traditional behavioral techniques do not allow for observations focusing on individuals [46]. Group size is important as smaller groups of laying hens (estimated at less than 70 individuals) are believed to adopt different social structures than those in larger groups (reviewed by [47]). In that sense, relationships identified in small groups may not be applicable to larger groups, and, by extension, the world's commercial laying hens that are typically housed in flocks containing 5,000 to 50,000 animals (when not housed in cages). In summary, the current experimental set up has unique features to allow the identification of relationships between behavioral patterns and brain DNA methylation.

The relationship between behavioural and epigenetic profiles

The results of the current study support the position that epigenetic variation within brain tissues is an essential factor to explain the natural emergence of behavioural variation. In our case, we focused on the nidopallium, which is a region of the avian brain involved in complex cognitive abilities, such as executive function, and considered to be analogous to the mammalian pre-frontal cortex [21].

By applying established classification schemes in a novel way to characterize animal movement and location patterns over the study period, we identified different and unique patterns across individuals. In comparison to the simple individual-level variables (e.g., time inside), we believe

complex metrics (i.e., DTW clustering, Entropy, and the Leading Index) offer a more comprehensive representation of an individual animal's behaviour.

Importantly, the different behavioural patterns that spontaneously emerged and were able to be documented in our population of hens were linked to related DNA methylation patterns in the nidopallium. Although a random sample of hens with the larger population would be expected to show some degree of epigenetic variation, the combination of novel movement classification analysis and epigenetic methods has revealed the emergence of linkages between the two that would have not been apparent otherwise. Unique DMRs were identified between behavioural patterns that emerged within all classification schemes, among which, entropy was associated with the greatest number (i.e., 107) (Figure 3A).

The biological mechanisms behind unique DMRs emerging between behavioral patterns is only conjecture at this point. However, it is important to consider that each classification scheme measures different properties of the time series. For instance, the type of entropy assessed in the current study may reflect activity levels. Entropy, which addressed the complexity of a given time series, was highly correlated with the number of transitions. This probably means that the kind of entropy measured here could somehow relate to increased exploratory behavior, which would be reflected in increased number of transitions between areas. In contrast, DTW clustering provides a more comprehensive representation of the time series as it incorporates both location as well as the timestamp into the resulting metric. This additional information could make DTW sensitive to biological processes that would influence the decision of animals of reaching particular areas, as well as the duration and time of day of their presence at these areas. For instance, increased time in areas external to the barn interior would result in proportionally greater exposure to sunlight, better air quality (particularly the free range and stone yard area), or risks associated with particular areas (e.g., use of the free range and stone yard area would have a greater chance of predation). The leading index was the only metric that represented an aspect of social behavior. Thus, it is interesting

to consider leading index in light of decision-making processes, reactivity, or behavioral tendencies towards conspecifics (e.g, aggression or boldness). Our results show that these properties described may have a functional relationship to the associated DMR.

Among the DMRs found, 12 were selected with a relatively stringent P value cut-off (P ≤0.005, Table 1). Of these DMRs, eight are gene related. Moreover, two of these DMRs have P values ≤ 0.0005 and relate to the genes Netrin G1 (DTW scheme) and Cadherin-22 (entropy scheme). We considered all these eight genes to be of special interest, and thus, their function is discussed here. Three of these genes exhibit an important role in brain function or development. Netrin G1 is part of a conserved family of proteins involved in axon guidance during developmental phases of the vertebrate nervous system [48]. Cadherin-22 is suggested to regulate cell-cell adhesion in morphogenesis and tissue formation in neural and non-neural cells in brain and neuroendocrine organs during developmental stages, as well as in the maintenance of these organs [49]. The *SLC9A9* gene, in turn, is a sodium hydrogen exchanger present in the recycling endosome, which is highly expressed in the brain and implicated in neuropsychiatric disorders such as autism spectrum disorders [50]. Other two genes in the list have important roles in the cell cycle checkpoints and exit. For example, the *ATR* gene encodes a serine/threonine kinase involved in damage-induced G2 checkpoint control and apoptosis in proliferating cells [51]. The *MST1* gene, in turn, is part of the Hippo signaling pathway, which in the optic neuroepithelia of Drosophila is shown to participate in cell cycle exit [52].

Association between classification schemes and genomic regulatory regions

A deeper exploration of the linkages between behavioural classification schemes and differential DMRs revealed that the schemes differentially associated with specific genomic regulatory regions. For instance, while similar proportion of DMRs associated with intronic, downstream, and exonic regions across the three schemes, DTW had double the proportion of affected promoter regions and half of the distal intergenic regions. DMRs exhibiting hypermethylation of promoter regions are

expected to inhibit the transcriptional machinery and thus prevent gene expression, according to standard assumptions [53]. Alternatively, promoter hypomethylation generally activates genes as a result of increased accessibility of DNA by polymerase [53]. In this sense, the DMRs linked with the differential behavioral clusters identified by DTW could have a direct role on the abundance associated with the expression of specific gene products. The effect of distal intergenic regions on gene expression is less clear and dependent on the region's association with other factors such as enhancers. Enhancers regulated by DNA methylation are linked to the coordinated transcriptional and epigenomic regulation of developmental genes in vertebrates [54].

It is interesting to consider the usefulness of DTW for obtaining clusters of individuals that are more likely to be exposed to particular environments or engage in particular activities. For example, Cluster 3 formed by DTW, with its limited exposure to natural sunlight, could relate to a specialized suite of gene products (e.g., vitamin D conversion) that would be categorically different from the gene products that could relate to activities of animals in the other clusters which routinely exit the barn interior. Alternatively, the association of the Leading Index classification scheme with nearly 70% of the distal intergenic regions would allow for a relatively greater flexibility in gene expression regulation, rather than simply regulating expression levels. This flexibility refers to the ability of one genomic region or element to regulate the expression of different genes, such as in epigenetic clusters of regulation. In this sense, the social behavior of the animals as characterized by the Leading Index will be highly dependent not only on their own behavior but also on the behavior of their conspecifics, which would, in turn, be related to a flexible ability of the genome to regulate gene expression. For instance, an animal classified as a follower in the Leading Index will enter into an area only if initiated by a specific conspecific in contrast to animals classified in the DTW Cluster 2 which reliably go to all areas of the pen moved by their own initiative.

Journal Pre-proof

DMR-associated Genes Enriched in Biological Pathways

Subsequent enrichment analysis of the genes associated to the DMRs revealed that some DMR-associated genes such as RAF1, GSL and TGFβ2 are influential on several important biological pathways. RAF1 is associated to a DMR that emerged in the comparison between DTW 2 and 1. RAF1 is an oncogene with suggested involvement in the pathogenesis of glioblastoma, the most aggressive type of brain cancer [55]. (RAF1 is member of the RAS/extracellular signal-regulated kinase 1/2 signalling pathway. Mutations in these genes are known to associate with so-called RASopathies, the most common of these being the Noonan Syndrome [56]. The effects of the Noonan Syndrome include structural malformations, developmental delays, but also behavioural issues such as irritability and communication difficulties [57]. Interestingly, research shows that mice expressing the Raf1L613V gain-of-function mutation, associated with the Noonan Syndrome, perform better than controls in some aspects of common behavioural tests such as the water radial-arm maze, Morris water maze, and cued fear conditioning tasks [56]. GSL, in turn, is a glutamine synthetase-like gene. In the brain, excesses of ammonia and the neurotransmitter glutamate are regulated by their conversion to glutamine. This happens mainly in astrocytes and by the action of the glutamine synthetase enzyme [58]. Higher glutamine levels in the prefrontal cortex of mammals are associated with better performance in a reversal learning task [59] and attenuation of depressive behaviour [60]. TGF Betas, in general, are molecular components of the signaling cascades defining the development and survival of many neuronal groups, and TGFβ2, in particular, is relatively more important during development [61]. Interestingly, TGFβ2 mutant mice exhibit less caudal 5-HT neurons and impaired development of raphe neurons during embryogenesis, as well as lower serotonin levels in the hindbrain and cortex in adulthood [61]. Another interesting result was the presence of three DMR-associated genes from the low vs high entropy comparison in the p53 signaling pathway. Described nearly 40 years ago and known as '*The Guardian of the Genome*', p53 is one of the most important and well-studied tumor suppressor factors [62]. The DMR-associated genes that participate in the p53 pathway are SHISA5, RFWD2 and ATR. SHISA5, also known as

SCOTIN in humans, is a pro-apoptotic gene induced after DNA damage or cellular stress in a p53-dependent manner [63]. No role has been reported for SHISA5 in relation to behavior. ATR (which is among our genes with the lowest p values, Table 1) is shown to maintain chromosomal integrity during postnatal cerebellar neurogenesis [64]. Interestingly, RFWD2, which is hypermethylated in the promoter regions of low entropy chickens (Suppl Table S5), is also hypermethylated in the prefrontal cortex (the mammalian equivalent to the bird nidopallium) of schizophrenia human patients [65]. Overall, our results suggest that behavior patterns share molecular mechanisms and are interrelated with other biological functions, such as tumorigenesis (related to the RAF1 oncogene) and cell cycle checkpoints and exit to repair (the main described functions of the p53 pathway). These genes and pathways they participate provide hints to explain the molecular basis of behavioral patterns.

Word clouds obtained from the pathways enriched by the gene associated DMRs give an idea of the different biological functions that could be related to each behavioural pattern identified in each classification scheme. Our data shows that while Entropy relates to cell cycle checkpoints, Leading Index relates to mitochondrial function, and DTW relates to regulation of gene expression. These results suggest that different behaviour patterns could be linked to different biological mechanisms within neurons. For example, according to our data acting as a leader or a follower (e.g. a classification involving variable decision making abilities) would be mostly associated with mitochondrial function and regulation of energy expenditure, while behavioural consistency (e.g., habits or repeated behaviours such as those observed within DTW clusters) would be related to fine-tuned regulation of gene expression at the level of transduction and activation/inactivation of proteins (e.g., by phosphorylation). These results open the possibilities for exciting experiments to investigate the biological basis of different behaviours. The methodologies described here could be employed for future investigations of causal relationships between epigenetic changes, cellular function and behavioral patterns.

Journal Pre-proof

While we cannot causally link the specific classification schemes and differentiated DMRs, the present paper shows a clear relationship between DNA methylation and specific behavioral patterns. The linkage between behavioural and epigenetic patterns is not in itself novel. For instance, post-traumatic stress disorder in humans is well documented to be correlated with epigenetic changes in the brain [8]. Behaviours linked to epigenetic patterns are also well documented in experimental paradigms where conditions or behaviours are experimentally induced in animal models, e.g. maternal separation in macaques [2] or maternal grooming in mice [4]. One of the novelties of our experiment lies in that the observed patterns (i.e., both behavioural and epigenetic) arose spontaneously within a group of individuals living in a relatively controlled and homogenous environment. The emergence of these patterns in animals that are able to 'freely' explore and engage in a full repertoire of behaviours [66] represents a unique experimental condition to understand factors influencing phenotypic variation in addition to genetic conformation. Although we are limited in our ability to draw conclusions about the nature of these relationships in terms of associated causal mechanisms, our work establishes an important foundation for future directed hypothesis-based evaluation of the link between neural epigenetics and behavioural variability in vertebrates. Our study suggests that different biological functions within neurons could be responsible for the emergence of different behaviours.

Journal Pre-proof

Acknowledgements

We greatly appreciate funding from the Haldimann and Eva Husi Foundations for financial support for renting the commercial barns, as well as technical support provided by Markus Schwab in organizing and installing the tracking equipment. We also greatly appreciate funding from the Swedish Research Council (FORMAS) grants #2018-01074 and #2017-00946 to CGB. We appreciate support by the COST Action STSM for KeelBoneDamageNet (CA15224) which funded a writing retreat that made it possible the writing of the bulk of this manuscript. F.P. appreciates funding from

FAPESP projects #2016/20440-3 and #2018/13600-0. The authors appreciate the support from the National Genomics Infrastructure in Stockholm funded by Science for Life Laboratory, the Knut and Alice Wallenberg Foundation and the Swedish Research Council. Likewise, the authors appreciate the computational resources provided by the SNIC/Uppsala Multidisciplinary Center for Advanced Computational Science, for assistance with massively parallel sequencing and for access to the UPPMAX computational infrastructure.

Data access

The dataset supporting the conclusions of this article is available in the European Nucleotide Archive (ENA) repository (EMBL-EBI), under accession PRJEB36004, which can be reached through the following link: http://www.ebi.ac.uk/ena/data/view/ PRJEB36004.

Figure Legends

Figure 1- Representative movement patterns of three clusters obtained through DTW. Each timeseries shown reflects the typical movement pattern of each cluster throughout one day in relation to the four distinct areas. Each line shows the position of the hen representing the median point obtained from the analysis of all the behavioral patterns recorded within a specific cluster. The four areas are: inside the barn (IN), winter garden (WG), stone yard (SY) and free range (FR).

Figure 2- Heat maps depicting DMRs (P>0.05) in the nidopallium of laying hens who were identified as belonging to different behavioral classification schemes: (A) Entropy, (B) Leading Index and DTW clusters (C-E). Animals with leading indices above the median were classified as "Leaders" while the others as "Followers". Entropy values above the median value are considered high, and those below the median, low. DTW clusters based on dissimilarity of time patterns in movement between areas were identified using hierarchical clustering analysis.

Figure 3- Venn diagrams showing the number of: (A) unique DMRs found within the different behavioral classification schemes, as well as overlapping DMRs among these schemes, and (B) DMRs found among different DTW clusters, as well as overlapping DMRs obtained after each pairwise comparison..

Figure 4- Pie charts representing genomic functional features mapped to DMRs ($P<0.05$) in relation to nearby chicken genes (based on the chicken reference genome) for each classification scheme (entropy, leading index and dynamic time warping).

Figure 5- (A) Extended gene networks obtained with genemania.org based on the genes related to DMRs found among the different behavioral classification schemes. (B) Word clouds were created from the Genemania functional analysis performed on each extended network obtained in (A). The word size is relative to how often the term was repeated in the output of the Genemania functional analysis.

Journal Pre-proof

References

[1] P. Jensen, Adding 'epi-' to behaviour genetics: implications for animal domestication, J Exp Biol 218(Pt 1) (2015) 32-40.
[2] N. Provencal, M.J. Suderman, C. Guillemin, R. Massart, A. Ruggiero, D. Wang, A.J. Bennett, P.J. Pierre, D.P. Friedman, S.M. Cote, M. Hallett, R.E. Tremblay, S.J. Suomi, M. Szyf, The signature of maternal rearing in the methylome in rhesus macaque prefrontal cortex and T cells, The Journal of neuroscience : the official journal of the Society for Neuroscience 32(44) (2012) 15626-42.
[3] D. Wang, M. Szyf, C. Benkelfat, N. Provencal, G. Turecki, D. Caramaschi, S.M. Cote, F. Vitaro, R.E. Tremblay, L. Booij, Peripheral SLC6A4 DNA methylation is associated with in vivo measures of human brain serotonin synthesis and childhood physical aggression, PloS one 7(6) (2012) e39501.
[4] I.C. Weaver, N. Cervoni, F.A. Champagne, A.C. D'Alessio, S. Sharma, J.R. Seckl, S. Dymov, M. Szyf, M.J. Meaney, Epigenetic programming by maternal behavior, Nature neuroscience 7(8) (2004) 847-54.
[5] M.K. Skinner, M. Manikkam, C. Guerrero-Bosagna, Epigenetic transgenerational actions of environmental factors in disease etiology, Trends Endocrinol Metab 21(4) (2010) 214-22.
[6] C. Guerrero-Bosagna, M.K. Skinner, Environmentally induced epigenetic transgenerational inheritance of phenotype and disease, Molecular and cellular endocrinology 354(1-2) (2012) 3-8.
[7] J. Graff, I.M. Mansuy, Epigenetic codes in cognition and behaviour, Behav Brain Res 192(1) (2008) 70-87.
[8] S. Malan-Muller, S. Seedat, S.M. Hemmings, Understanding posttraumatic stress disorder: insights from the methylome, Genes Brain Behav 13(1) (2014) 52-68.
[9] S.D. Gosling, From mice to men: what can we learn about personality from animal research?, Psychol Bull 127(1) (2001) 45-86.
[10] D. Reale, N.J. Dingemanse, A.J. Kazem, J. Wright, Evolutionary and ecological approaches to the study of personality, Philos Trans R Soc Lond B Biol Sci 365(1560) (2010) 3937-46.
[11] C. Rufener, J. Berezowski, F. Maximiano Sousa, Y. Abreu, L. Asher, M.J. Toscano, Finding hens in a haystack: Consistency of movement patterns within and across individual laying hens maintained in large groups, Sci Rep 8(1) (2018) 12303.
[12] M. Gerken, J. Petersen, Heritabilities for Behavioral and Production Traits in Japanese Quail (Coturnix coturnix japonica) Bidirectionally Selected for Dustbathing Activity, Poultry Science 71(5) (1992) 779-788.
[13] L. Hradecká, L. Bartoš, I. Svobodová, J. Sales, Heritability of behavioural traits in domestic dogs: A meta-analysis, Appl. Anim. Behav. Sci. 170 (2015) 1-13.
[14] U. Konig von Borstel, B. Tonepohl, A.K. Appel, B. Voss, H. Brandt, S. Naderi, M. Gauly, Suitability of traits related to aggression and handleability for integration into pig breeding programmes: Genetic parameters and comparison between Gaussian and binary trait specifications, PloS one 13(12) (2018) e0204211.
[15] M.J. Wade, P. Bijma, E.D. Ellen, W. Muir, Group selection and social evolution in domesticated animals, Evol Appl 3(5-6) (2010) 453-65.
[16] O. Tal, E. Kisdi, E. Jablonka, Epigenetic contribution to covariance between relatives, Genetics 184(4) (2010) 1037-50.
[17] L. Varona, S. Munilla, E.F. Mouresan, A. Gonzalez-Rodriguez, C. Moreno, J. Altarriba, A Bayesian model for the analysis of transgenerational epigenetic variation, G3 5(4) (2015) 477-85.
[18] F. Johannes, M. Colome-Tatche, Quantitative epigenetics through epigenomic perturbation of isogenic lines, Genetics 188(1) (2011) 215-27.
[19] R. Singal, G.D. Ginder, DNA methylation, Blood 93(12) (1999) 4059-70.
[20] L.F. Rizzardi, P.F. Hickey, V. Rodriguez DiBlasi, R. Tryggvadottir, C.M. Callahan, A. Idrizi, K.D. Hansen, A.P. Feinberg, Neuronal brain-region-specific DNA methylation and chromatin accessibility are associated with neuropsychiatric trait heritability, Nature neuroscience 22(2) (2019) 307-316.

Journal Pre-proof

[21] C. Herold, N. Palomero-Gallagher, B. Hellmann, S. Kroner, C. Theiss, O. Gunturkun, K. Zilles, The receptor architecture of the pigeons' nidopallium caudolaterale: an avian analogue to the mammalian prefrontal cortex, Brain Struct Funct 216(3) (2011) 239-54.
[22] Y. Gómez, J. Berezowski, Y. Abreu, F.M. Sousa, S.G. Gebhardt-Henrich, M.J. Toscano, A. Stratmann, S. Vögeli, B. Voelkl, Ranging patterns and social associations in laying hens, 53rd Congress of the International Society for Applied Ethology (ISAE), Bergen, Norway, 2019.
[23] S.G. Gebhardt-Henrich, M.J. Toscano, E.K.F. Fröhlich, Use of outdoor ranges by laying hens in different sized flocks, Appl. Anim. Behav. Sci. 155 (2014) 74-81.
[24] S.G. Gebhardt-Henrich, E.K.F. Fröhlich, F. Burose, J. Fleurent, M. Gantner, M. Zähner, Individual tracking of laying hens with an RFID-System (Registrierung des Auslaufverhaltens einzelner Legehennen mit einem RFID-System), Landtechnik 69(9) (2014) 301-307.
[25] J.S. Richman, J.R. Moorman, Physiological time-series analysis using approximate entropy and sample entropy, Am J Physiol Heart Circ Physiol 278(6) (2000) H2039-49.
[26] R. Bellman, R. Kalaba, On adaptive control processes, IRE T Automat Contr 4(2) (1959) 1-9.
[27] E. Keogh, C.A. Ratanamahatana, Exact indexing of dynamic time warping, Knowl Inf Syst 7(3) (2005) 358-386.
[28] A. Sardá-Espinosa, Time-Series Clustering in R Using the dtwclust Package, The R Journal 11(1) (2019).
[29] L. Pulles, The chick brain in stereotaxic coordinates : An atlas featuring neuromeric subdivisions and mammalian homologies, 1st ed., Academic Press2007.
[30] F. Pertille, C. Guerrero-Bosagna, V.H. Silva, C. Boschiero, R. Nunes Jde, M.C. Ledur, P. Jensen, L.L. Coutinho, High-throughput and Cost-effective Chicken Genotyping Using Next-Generation Sequencing, Sci Rep 6 (2016) 26929.
[31] C. Guerrero-Bosagna, P. Jensen, Optimized method for methylated DNA immuno-precipitation, MethodsX 2 (2015) 432-9.
[32] F. Pertille, M. Brantsaeter, J. Nordgreen, L.L. Coutinho, A.M. Janczak, P. Jensen, C. Guerrero-Bosagna, DNA methylation profiles in red blood cells of adult hens correlate with their rearing conditions, J Exp Biol 220(Pt 19) (2017) 3579-3587.
[33] R.J. Elshire, J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, E.S. Buckler, S.E. Mitchell, A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species, PloS one 6(5) (2011) e19379.
[34] J.A. Poland, P.J. Brown, M.E. Sorrells, J.L. Jannink, Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach, PloS one 7(2) (2012) e32253.
[35] S. Andrews, FastQC: a quality control tool for high throughput sequence data, Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc, 2010.
[36] B. Langmead, S.L. Salzberg, Fast gapped-read alignment with Bowtie 2, Nature methods 9(4) (2012) 357-9.
[37] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, S. Genome Project Data Processing, The Sequence Alignment/Map format and SAMtools, Bioinformatics 25(16) (2009) 2078-9.
[38] J.M. Catchen, A. Amores, P. Hohenlohe, W. Cresko, J.H. Postlethwait, Stacks: building and genotyping Loci de novo from short-read sequences, G3 (Bethesda) 1(3) (2011) 171-82.
[39] M. Lienhard, C. Grimm, M. Morkel, R. Herwig, L. Chavez, MEDIPS: genome-wide differential coverage analysis of sequencing data derived from DNA enrichment experiments, Bioinformatics 30(2) (2014) 284-6.
[40] M.D. Robinson, A. Oshlack, A scaling normalization method for differential expression analysis of RNA-seq data, Genome Biol 11(3) (2010) R25.
[41] Y. Zhang, T. Liu, C.A. Meyer, J. Eeckhoute, D.S. Johnson, B.E. Bernstein, C. Nusbaum, R.M. Myers, M. Brown, W. Li, X.S. Liu, Model-based analysis of ChIP-Seq (MACS), Genome Biol 9(9) (2008) R137.
[42] G. Yu, L.G. Wang, Q.Y. He, ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization, Bioinformatics 31(14) (2015) 2382-3.

[43] D. Warde-Farley, S.L. Donaldson, O. Comes, K. Zuberi, R. Badrawi, P. Chao, M. Franz, C. Grouios, F. Kazi, C.T. Lopes, A. Maitland, S. Mostafavi, J. Montojo, Q. Shao, G. Wright, G.D. Bader, Q. Morris, The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function, Nucleic acids research 38(Web Server issue) (2010) W214-20.
[44] G.G. Mittelbach, N.G. Ballew, M.K. Kjelvik, Fish behavioral types and their ecological consequences, Can. J. Fish. Aquat. Sci. 71 (2014) 927-944.
[45] S.H. Richter, S. Hintze, From the individual to the population–and back again? Emphasising the role of the individual in animal welfare science, Appl. Anim. Behav. Sci. 212 (2019) 1-8.
[46] J.M. Siegford, J. Berezowski, S.K. Biswas, C.L. Daigle, S.G. Gebhardt-Henrich, C.E. Hernandez, S. Thurner, M.J. Toscano, Assessing Activity and Location of Individual Laying Hens in Large Groups Using Modern Technology, Animals (Basel) 6(2) (2016).
[47] C.C. Croney, R.C. Newberry, Group size and cognitive processes, Appl. Anim. Behav. Sci. 103 (2007) 215-228.
[48] T. Nakashiba, T. Ikeda, S. Nishimura, K. Tashiro, T. Honjo, J.G. Culotti, S. Itohara, Netrin-G1: a novel glycosyl phosphatidylinositol-linked mammalian netrin that is functionally divergent from classical netrins, The Journal of neuroscience : the official journal of the Society for Neuroscience 20(17) (2000) 6540-50.
[49] K. Sugimoto, S. Honda, T. Yamamoto, T. Ueki, M. Monden, A. Kaji, K. Matsumoto, T. Nakamura, Molecular cloning and characterization of a newly identified member of the cadherin family, PB-cadherin, J Biol Chem 271(19) (1996) 11548-56.
[50] J. Patak, J.L. Hess, Y. Zhang-James, S.J. Glatt, S.V. Faraone, SLC9A9 Co-expression modules in autism-associated brain regions, Autism Res 10(3) (2017) 414-429.
[51] V. Enriquez-Rios, L.C. Dumitrache, S.M. Downing, Y. Li, E.J. Brown, H.R. Russell, P.J. McKinnon, DNA-PKcs, ATM, and ATR Interplay Maintains Genome Integrity during Neurogenesis, The Journal of neuroscience : the official journal of the Society for Neuroscience 37(4) (2017) 893-905.
[52] B.V. Reddy, C. Rauskolb, K.D. Irvine, Influence of fat-hippo and notch signaling on the proliferation and differentiation of Drosophila optic neuroepithelia, Development 137(14) (2010) 2397-408.
[53] D. Anastasiadi, A. Esteve-Codina, F. Piferrer, Consistent inverse correlation between DNA methylation of the first intron and gene expression across tissues and species, Epigenetics Chromatin 11(1) (2018) 37.
[54] O. Bogdanovic, A.H. Smits, E. de la Calle Mustienes, J.J. Tena, E. Ford, R. Williams, U. Senanayake, M.D. Schultz, S. Hontelez, I. van Kruijsbergen, T. Rayon, F. Gnerlich, T. Carell, G.J. Veenstra, M. Manzanares, T. Sauka-Spengler, J.R. Ecker, M. Vermeulen, J.L. Gomez-Skarmeta, R. Lister, Active DNA demethylation at enhancers during the vertebrate phylotypic period, Nature genetics 48(4) (2016) 417-26.
[55] Z. Liu, Y. Liu, L. Li, Z. Xu, B. Bi, Y. Wang, J.Y. Li, MiR-7-5p is frequently downregulated in glioblastoma microvasculature and inhibits vascular endothelial cell proliferation by targeting RAF1, Tumour Biol 35(10) (2014) 10177-84.
[56] M.C. Holter, L.T. Hewitt, S.V. Koebele, J.M. Judd, L. Xing, H.A. Bimonte-Nelson, C.D. Conrad, T. Araki, B.G. Neel, W.D. Snider, J.M. Newbern, The Noonan Syndrome-linked Raf1L613V mutation drives increased glial number in the mouse cortex and enhanced learning, PLoS genetics 15(4) (2019) e1008108.
[57] A. Wood, A. Massarano, M. Super, R. Harrington, Behavioural aspects and psychiatric findings in Noonan's syndrome, Arch Dis Child 72(2) (1995) 153-5.
[58] I. Suarez, G. Bodega, B. Fernandez, Glutamine synthetase in brain: effect of ammonia, Neurochem Int 41(2-3) (2002) 123-42.
[59] A. Lacreuse, C.M. Moore, M. LaClair, L. Payne, J.A. King, Glutamine/glutamate (Glx) concentration in prefrontal cortex predicts reversal learning performance in the marmoset, Behav Brain Res 346 (2018) 11-15.
[60] H. Son, J.H. Baek, B.S. Go, D.H. Jung, S.B. Sontakke, H.J. Chung, D.H. Lee, G.S. Roh, S.S. Kang, G.J. Cho, W.S. Choi, D.K. Lee, H.J. Kim, Glutamine has antidepressive effects through increments of

glutamate and glutamine levels and glutamatergic activity in the medial prefrontal cortex, Neuropharmacology 143 (2018) 143-152.
[61] E. Chleilat, R. Mallmann, R. Spanagel, N. Klugbauer, K. Krieglstein, E. Roussa, Spatiotemporal Role of Transforming Growth Factor Beta 2 in Developing and Mature Mouse Hindbrain Serotonergic Neurons, Front Cell Neurosci 13 (2019) 427.
[62] A.B. DeLeo, E. Appella, The p53 Saga: Early Steps in the Development of Tumor Immunotherapy, J Immunol 204(9) (2020) 2321-2328.
[63] R.K. Gupta, R. Tripathi, B.J. Naidu, U.K. Srinivas, L.S. Shashidhara, Cell cycle regulation by the pro-apoptotic gene Scotin, Cell Cycle 7(15) (2008) 2401-8.
[64] P.Y. Lang, G.J. Nanjangud, M. Sokolsky-Papkov, C. Shaw, D. Hwang, J.S. Parker, A.V. Kabanov, T.R. Gershon, ATR maintains chromosomal integrity during postnatal cerebellar neurogenesis and is required for medulloblastoma formation, Development 143(21) (2016) 4038-4052.
[65] S.A. Lee, K.C. Huang, Epigenetic profiling of human brain differential DNA methylation networks in schizophrenia, BMC Med Genomics 9(Suppl 3) (2016) 68.
[66] C.A. Weeks, C.J. Nicol, Behavioural needs, priorities and preferences of laying hens, Worlds Poult Sci J 62(2) (2006) 296-307.

Journal Pre-proof

Genomic location	Width (bp)	CpGs	Annotation	Gene Id	Entrez ID	Gene Symbol	Ger
chr26:870542-871014	472	1	Intron	ENSGALG00000031122.1	424336	NTNG1	Net
chr4:10886922-10887074	152	1	Promoter	ENSGALG00000007211.5	419305	CDH22	Cad
chr20:10877159-10877501	342	14	Intron	ENSGALG00000002374.5	424772	SLC9A9	Solu 9 m
chr2:82081027-82081519	492	11	Distal Intergenic	ENSGALG00000040186.2	NA	NA	NA
chr19:7725868-7726161	293	21	Distal Intergenic	ENSGALG00000002892.5	423653	SLC22A15L	Solu 22 r
chr15:7902061-7902648	587	27	Distal Intergenic	ENSGALG00000005415.5	403089	TEAD1	TEA tran 1
chr20:10489420-10489893	473	17	Exon	ENSGALG00000002663.5	424777	ATR	ATR seri kina
chr12:2458270-2458563	293	17	Promoter	ENSGALG00000002722.5	396135	MST1	Ma stin (he fact
chrZ:73604888-73605174	286	10	Intron	ENSGALG00000029598.2	NA	NA	NA
chr12:7918429-7918698	269	23	Distal Intergenic	ENSGALG00000041167.2	NA	NA	NA
chr24:2807083-2807962	879	12	Promoter	ENSGALG00000043955.1	NA	NA	NA
chr10:20047539-20048245	706	39	Promoter	ENSGALG00000036163.1	100858113	ZNF710	Zinc 710

Journal Pre-proof

Comparisson	DMR-Associated Gene enriched in Pathways	Pathways involved
DTW cluster 2 vs 1	ABCB6	ABC transporters
	FANCE	Fanconi anemia
	RAF1	VEGF signaling; VEGF signalling; ErbB signalling; Progesterone-mediated oocyte maturation; Gap junction; GnRH signaling; C-type lectin receptor signaling; Melanogenesis; Vascular smooth muscle contraction; Apelin signaling; FoxO signaling; Autophagy - animal; Apoptosis; Influenza A; mTOR signaling; Cellular senescence; Focal adhesion; Regulation of actin cytoskeleton
DTW cluster 3 vs 2	NTNG1	Cell adhesion molecules (CAMs)
	ATP2B4	Adrenergic signaling in cardiomyocytes; Calcium signaling
DTW cluster 3 vs 1	GSL	Arginine biosynthesis; Nitrogen metabolism; Glyoxylate and dicarboxylate metabolism; Alanine, aspartate and glutamate metabolism; Biosynthesis of amino acids; Necroptosis
Entropy: Low vs High	RFWD2	p53 signaling
	ATR	
	SHISA5	
Leading index: Follower vs Leader	TGFB2	TGF-beta signaling; AGE-RAGE signaling pathway in diabetic complications; Cell cycle; FoxO signaling; Cellular senescence
	AIFM1	Apoptosis; Necrosis
	ATP6V1AL	Phagosome; mTOR signaling

Journal Pre-proof

Conflict of interest declaration:

The authors declare no conflict of interests.

Journal Pre-proof

Highlights of the paper:

'DNA methylation variation in the brain of laying hens in relation to differential behavioral patterns'

- Behavioral patterns that spontaneously emerge in hens living in a highly controlled environment were identified with a unique tracking system.
- Behavioral activity patterns were characterized through three classification schemes.
- Differentially methylated regions (DMRs) were identified in the nidopallium between behavioural patterns emerging within classification schemes.
- DMRs from the classification schemes associated with different genomic features.
- DMR-associated genes within each classification schemes had distinctive biological functions.

Journal Pre-proof