INTRODUCTION

Mechanisms that conserve water are a necessary prerequisite for terrestrial life. Optimal hydration is essential for survival, and evolutionarily conserved neuroendocrine mechanisms ensure that water is not lost, especially following periods of dehydration.\(^1,2\) The brain circuits regulating water balance are centred on the magnocellular neurones (MCNs) of the supraoptic nucleus (SON). These circuits are responsible for the production and release of antidiuretic hormone (ADH) and oxytocin, which control water reabsorption and lactation, respectively.

We have used RNA sequencing to describe the polyadenylated transcriptome of the SON of the male Wistar Han rat. These data have been mined to generate comprehensive catalogues of functional classes of genes (enzymes, transcription factors, endogenous peptides, G protein coupled receptors, transporters, catalytic receptors, channels and other pharmacological targets) expressed in this nucleus in the euhydrated state, and that together form the basal substrate for its physiological interactions. We have gone on to show that fluid deprivation for 3 days (dehydration) results in changes in the expression levels of 2247 RNA transcripts, which have similarly been functionally catalogued, and further mined to describe enriched gene categories and putative regulatory networks (Regulons) that may have physiological importance in SON function related plasticity.

We hope that the revelation of these genes, pathways and networks, most of which have no characterised roles in the SON, will encourage the neuroendocrine community to pursue new investigations into the new ‘known-unknowns’ reported in the present study.

KEYWORDS

dehydration, plasticity, regulon, RNA sequencing, supraoptic nucleus, transcriptome
large peptidergic neurones project axons via the internal zone of the median eminence to axon terminals located adjacent to blood capillaries in the pars nervosa (PN) neurovascular interface of the pituitary gland. Two discrete populations of SON MCNs produce two major neuropeptide hormones secretory products, namely arginine vasopressin (AVP, encoded by the Avp gene) and oxytocin (OXT, encoded by the Oxt gene). The mature peptide hormone processing products of these genes are stored in PN nerve terminals until secreted into the systemic circulation in response to appropriate physiological cues. Dehydration results in a rise in plasma osmolality that is detected by intrinsic MCN osmosensory mechanisms, and by osmoreceptive neurones in the circumventricular organs such as the subfornical organ which provide excitatory inputs to MCNs. Both of these processes alter the firing of AVP MCNs leading to hormone release. AVP travels through the blood stream to specific receptors located in the kidney collecting duct, where it promotes water reabsorption, and the thick ascending limb of the loop of Henle, where it elicits sodium reabsorption. As a consequence of the secretion of AVP during dehydration, there is a need to replenish PN stores. Thisstarts with an up-regulation in Avp gene transcription, which results in an increase in the abundance of the precursor heteronuclear RNA, and the mature mRNA, in addition to its well-established role in parturition and lactation. OXT is assumed to have natriuretic activity at the level of the kidney.

Chronic osmotic stimulation evokes a dramatic functional remodelling of the rat SON that affects both neurones and glia. A plethora of activity-dependent changes in the morphology, electrical properties and biosynthetic and secretory activity of the SON have been described, which contribute to the facilitation of hormone production and delivery, and hence survival. We, and others, have sought to understand this function-related plasticity at the level of the expression of the genome. This strategy involves subjecting rats to physiological cues, then carrying out transcriptomic analysis of specific brain regions. These data are then subjected to bioinformatic analysis to identify nodal target genes, the functions of which are then investigated in vivo. Thus, we used GeneChip™ (Affymetrix, Santa Clara, CA, USA) analysis to document the transcriptomes of the adult male and female Sprague-Dawley rat SON and PVN, and to describe how these change following the chronic osmotic challenge of 72 hours of complete fluid deprivation (dehydration). Some of the novel differentially expressed genes (DEGs) identified have been subjected to detailed functional investigation, for example Creb3l1, Azin1, Slc12a1, Caprin2, and Rad1. We have thus demonstrated that the SON is a tractable model that enables sustained progress from transcriptome datasets to novel physiological understanding. Note that gene symbols are used throughout this manuscript; the corresponding gene descriptions are provided in the Supporting information (Table S1).

Our 2006 Affymetrix analysis utilised the Rat 230 2.0 array, which enabled the simultaneous analysis of 31 099 probe sets, corresponding to 28 700 genes. Although robust and extensive, as well as state-of-the art at the time, these studies were limited with respect to gene number and annotation. Over the subsequent years, transcriptome technologies have become truly comprehensive with the development of RNA-sequencing (RNA-seq). Here, we employ next generation sequencing to describe the polyadenylated transcriptomes of the Wistar Han rat SON from 3-month-old euhydrated and dehydrated rats. We have mined these datasets to catalogue the basal expression profiles of categories of functionally relevant genes within the SON. We have gone on to describe how the transcriptome profile changes with dehydration. These data have been subject to detailed and robust analyses that revealed enriched functional classes of DEGs, and putative regulatory networks (Regulons). Comparison of the Wistar RNA-seq data with our Sprague-Dawley microarray data revealed a core set of common responsive genes that are presumably crucial for the orchestration of SON function-related plasticity in the rat.

## 2. MATERIALS AND METHODS

### 2.1 Animals

All experiments were performed under a Home Office UK licence held under, and in strict accordance with, the provisions of the UK Animals (Scientific Procedures) Act (1986); they had also been approved by the University of Bristol Animal Welfare and Ethical Review Board. We choose to use male Wistar Han rats from the international standardisation programme (IGS) in our study (Charles River, Portishead, England). This carefully managed breeding program minimises the impact of genetic drift so that colonies bred in different locations around the world are not significantly divergent from each other giving a level of continuity and reproducibility in studies performed in laboratories worldwide. Twelve rats aged 11 weeks were purchased from Charles River for this study. Rats were housed in groups of three for 2 weeks under a 14:10 hour light/dark photocycle (lights on 5.00 AM) at 22°C and 50%-60% relative humidity (v/v) with food and water available ad lib. Cages contained sawdust, bedding material and cardboard tubing for enrichment.

Rat cages were randomly assigned to two groups: control (free access to drinking water) and dehydrated (removal of drinking water for 72 hours). All rats were humanely killed by striking of the cranium (stunning) and then immediately decapitated with a small animal guillotine (Harvard Apparatus, Holliston, MA, USA). Trunk blood was collected in heparin-coated tubes and centrifuged at 1600 g for 15 minutes at 4°C. Plasma osmolality was measured by freezing point depression using a Roebeling micro-osmometer (Camlab, Over, UK). Plasma for AVP measures was collected in 1-mL aliquots and snap frozen in liquid nitrogen before being stored at −80°C. Brains were rapidly removed from the cranium and placed into a chilled rat brain matrix for separation of the forebrain from the hindbrain. The forebrain was placed cut edge down onto aluminium foil resting on pellets of dry ice and covered with powdered dry ice (within 3 minutes of stunning). Animal experiments were performed between 10.00 AM and midday.
2.2 | Plasma AVP measures

Trunk blood was extracted from 1 mL of plasma. Two sample volumes of ice-cold acetone were added and samples were vortexed for 1 minute. Protein precipitates were removed by centrifugation at 2500 g for 25 minutes at 4°C. The supernatant was transferred to a new tube and mixed with 2 mL of cold petroleum ether by vortexing for 1 minute. The tubes were left to stand for 1 minute at room temperature before discarding the upper phase. The lower phase solution was lyophilised using a freeze dryer (Benchtop Pro; Biopharma, Winchester, UK). Plasma AVP concentrations were determined by specific radioimmunoassay. 34

2.3 | RNA-seq

Bilateral punches of the SON were collected from 12 coronal slices with a 1 mm sample corer (Fine Scientific Tools, Heidelberg, Germany) using the optic chiasm as a reference as described. 21 The microtubes containing SON punches were maintained on dry ice prior to re-suspension by continuous vortexing for 1 minute in 400 µL of QIAGEN lysis reagent (Qiagen, Valencia, CA, USA). After a 10-minute incubation at room temperature, debris was pelleted by centrifugation at 12 000 g for 3 minutes. Then, 350 µL was removed and mixed with an equal volume of absolute ethanol. This mix was applied to a Zymo-Spin IIIC column and total RNA was extracted using a Direct-zol RNA MiniPrep extraction kit (Zymo Research, Irvine, CA, USA). Total RNA was eluted in a volume of 25 µL. RNA went through rigorous quality control checks to assess purity and integrity (Agilent BioAnalyzer; RNA TapeStation; Agilent, Santa Clara, CA, USA). The average RIN value was 8.5 (range 8.4-8.6) (see Supporting information, Table S2). Poly(A) enriched bulk RNA-seq libraries were constructed using Illumina TruSeq Stranded mRNA kits (Source Bioscience, Nottingham, UK). Libraries were loaded onto lanes of an Illumina NextSeq flowcell and sequenced using 75bp paired-end (PE) runs (Source Bioscience). Each sample generated > 35 million PE reads.

2.4 | RNA-seq data mining

RNA-seq alignment and subsequent data analysis were all performed in house using our high-performance computer; “Hydra” (PowerEdgeR820 12 core supercomputer equipped with 512 GB RAM and 12 x 1 TB HDD; Dell, Round Rock, TX, USA).

RNA-seq reads were processed using RTA and CASAVA from Illumina’s suite of sequencing software. This produced a series of compressed FASTQ files per library. All raw reads were merged per library, and pre-processed for quality assessment, adaptor removal, quality trimming and size selection using the FastQC toolkit to generate quality plots for all read libraries. 35 We adopted a phred30 quality cutoff (99.9% base call accuracy). Our pipeline accepts RNA-seq post-trimmed data as input, before ultimately producing output tables of differentially expressed transcripts. Paired-end (2 × 75 bp) raw input data is initially aligned with STAR to the sixth iteration of the Rattus norvegicus reference genome (Rn6). FeatureCounts is used to generate read counts for each gene present in the Rn6 genome annotation.37 Our pipeline then uses the DEseq2 (v1.28.1) package in R to call DEGs. 38 This allows us to predict DEGs with high confidence, and to utilise the predictions with low P-values in our downstream validation. Benjamini-Hochberg correction was used for multiple comparison. The analysis was sufficiently powered (n = 5) to reduce the false discovery rate, and to enable systems level analysis. 39 DEGs with P adjusted (PAdj) values of < 0.05 are considered significant.

ENSEMBL 40 was used to retrieve genome annotations using biomaRt (version 2.44.4); 41,42 Bioconductor (version 3.11.4)42,43 and AnnotDbi (version 1.50.3) packages in R. Regularised log transformation was applied for principal component analysis and heatmap plot generation. Scaled Venn diagrams were generated using VennDiagram (version 1.6.20) package in R. 44 Gene Ontology (GO) 45 and Kyoto Encyclopaedia of Genes and Genomes (KEGG)46,47 analysis was performed using ClusterProfiler (version 3.16.1) package in R. 48 Reactome pathway analysis was performed using ReactomePA (version 1.32.0) implementation in ClusterProfiler. 49 Over-representation analysis was carried by filtering genes for average expression in normalised reads across all samples (baseMean) > 10, PAdj < 0.05 and log2 fold-change (log2FC) = 0.5. All expressed genes filtered for baseMean > 10 were used as the background in the over-representation analysis and as the input gene list in the gene-set enrichment analysis (GSEA). 50 Benjamini-Hochberg correction (PAdj < 0.05) was used for multiple comparison correction for the pathway analysis.

Transcriptional regulatory network (Regulon) analysis was carried out using RTN (version 2.12.1) package in R. 52,53 A list of transcription factors screened for regulatory network enrichment in the SON was generated by annotating all expressed genes for associated GO terms and filtering the unique results for matching ‘DNA-binding transcription factor activity’ (GO:0003700) term. Additionally, a list of validated human transcription factor orthologue genes was also screened for regulon enrichment. This has led to the identification of a single additional regulon (Rbck1) in the Wistar SON.

All derived gene expression data was visualised using custom scripts written in R. 55

2.5 | Microarray data mining and comparison with RNA-seq datasets

Previously published microarray data (17; NCI PMDI: 16432224) was remixed with updated annotations using the NCBI GEO2R tool, which relies upon GEOquery (version 2/40.0)56 and Limma (version 2.26.8)57 packages within the R environment (version 3.2.3). This analysis provides a list of dehydration induced log2-FC differentially
regulated genes in the rat which satisfy a corrected $P$-value threshold $< 0.05$. For RNA-seq and microarray comparison, Affymetrix probe IDs (affy_rat230_2) were converted to matching Ensemble gene IDs using BiomaRt.\textsuperscript{41,42}

### RESULTS AND DISCUSSION

#### 3.1 | Experimental protocol

Our experimental protocol is summarised in Figure 1. Euhydrated (control) rats ($n = 5$) were compared with rats subjected to 72 hours of complete fluid deprivation (dehydration) ($n = 5$). As expected, dehydration resulted in a significant decrease in body weight and a significant increase in plasma osmolality and plasma AVP content (see Supporting information, Table S2). We sequenced the polyadenylated SON transcriptomes from these animals. Reads were processed using a bespoke RNA-seq pipeline as described in detail in the Materials and methods. A complete catalogue of expressed genes and DEGs can be found in the Supporting information (Table S3). All data have been banked at the NCBI Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE175461).

#### 3.2 | The basal euhydrated transcriptome of the rat SON

In total, 22 219 genes are predicted to be expressed in the SON region (no base mean cut-off). The basal transcriptome of the euhydrated SON region was explored using a subset of 15 277 genes with a base mean expression level $> 10$. The top 100 most highly expressed genes in the euhydrated SON region are shown in Figure 2. Transcripts encoded by the arginine vasopressin ($\text{Avp}$; top 10) and oxytocin ($\text{Oxt}$; top 17) genes were among the most highly expressed transcripts in our dataset indicating a highly precise punching of the hypothalamic SON used for RNA extractions. The 15 277 genes expressed in the euhydrated SON (normalised reads...
FIGURE 2  The top 100 most abundantly expressed genes in the supraoptic nucleus (SON) under control euhydrated conditions. Genes are listed by the average normalised uniquely-mapped read count in control (euhydrated) animals. The colour scale indicates differentially expressed genes (DEGs) in dehydration ($P_{\text{Adj}} < 0.05$); grey bars represent genes that were not differentially expressed ($P_{\text{Adj}} > 0.05$).
≥ 10) were catalogued on the basis of their identity as transcription factors or their physiological or pharmacological classifications (B-H, IUPHAR). Thus, we identified 923 transcription factors (see Supporting information, Figure S1), 957 enzymes (see Supporting information, Figures S2 and S3), 228 endogenous peptides (see Supporting information, Figure S4), 247 G protein coupled receptors (GPCRs) (see Supporting information, Figure S5), 390 transporters (see Supporting information, Figure S6), 200 catalytic receptors (see Supporting information, Figure S7), 207 channels (see Supporting information, Figure S8) and 144 other pharmacological targets (see Supporting information, Figure S9). Figure 3 summarises the top 5% most highly expressed genes in each of these different categories.

3.3 | Transcriptome dynamics in the SON following dehydration

Differential expression analysis using DEseq2 identified a total of 2247 DEGs, 924 of which were down-regulated and 1323 were up-regulated (Figure 4A). Notably, the majority of genes (19 912) with non-zero expression values did not breach the statistical significance threshold (Benjamini-Hochberg correction, \( P_{\text{Adj}} < 0.05 \)) (Figure 4A).

Principal component analysis using all genes revealed distinct separation between the transcriptomes of control and dehydrated samples. Principal component 1 (PC1) explained 51% of total variance between the samples and was attributable to the experimental condition (Figure 4B). Samples separated based on experimental condition following hierarchical clustering of all identified DEGs, highlighting a distinct expression phenotype in dehydration (Figure 4C).

Figure 5A presents the top 100 most significant DEGs (\( P_{\text{Adj}} < 5.01 \times 10^{-15} \)) ordered by fold-change in dehydration (relative expression). Next, we selected all DEGs (\( P_{\text{Adj}} < 0.005, n = 2247 \)) and highlighted the top 100 genes with the greatest expression fold-change in dehydration (relative expression) (Figure 5B). This highlighted significant genes with the greatest response to dehydration disregarding the underlying variation between the samples. We observed that log2FC (relative expression in dehydration compared to the control) tends to favour lowly expressed transcripts over the abundant genes (Figure 5B). To account for this, we selected the top 100 genes with the greatest expression change favouring highly expressed genes. Genes were filtered for differential expression (\( P_{\text{Adj}} < 0.005 \)) and the difference between the average normalised read counts between the control and dehydrated conditions (delta) was calculated. The top 100 genes with the greatest difference (delta) in dehydration are shown in Figure 5C and listed in the Supporting information (Table S4). This comparison is inherently biased towards highly expressed transcripts and represents expression change measured by absolute units instead of relative expression compared to the control condition. In this analysis Avp (log2FC = 0.46, \( P_{\text{Adj}} = 1.49 \times 10^{-2} \)) had the greatest positive change value. Another gene well-known to change expression in dehydration, Atf4 (log2FC = 0.52, \( P_{\text{Adj}} = 9.73 \times 10^{-49} \)) ranked as top 40 in this list despite modest log2FC in dehydration. Only nine genes intersected the top 100 DEGs listed by greatest log2FC and Delta: Gpd1, Vgf, Scg2, Th, Creb3l1, Oacyl, Pdyn, Arhgdib and Caprin2 (Figure 5C). These represent abundantly expressed genes in the hypothalamic SON that undergo the greatest expression change in response to 72 hours of dehydration in the Wistar rat.

We then catalogued the SON dehydration DEGs on the basis of their identity as transcription factors or their physiological or pharmacological classifications (B-H, IUPHAR) (Figure 6). Using a fold change cut-off of > 25%, we identified 49 transcription factors (Figure 6A), 38 enzymes (Figure 6B), 23 endogenous peptides (Figure 6C), 19 GPCRs (Figure 6D), 23 transporters (Figure 6E), 20 channels (Figure 6F) 13 catalytic receptors (Figure 6G) and six additional pharmacological targets (Figure 6H). These genes are promising candidates mediating the function related plasticity of the SON.

3.4 | GO and pathway analysis of SON dehydration DEGs

To identify gene categories, pathways and networks that might be functionally involved in the osmoregulatory plasticity of the SON, we performed pathway analysis on DEGs identified in response to dehydration using GO, KEGG and Reactome databases.

Over-representation analysis (ORA) identified 23 enriched terms in the GO:Molecular Function (GO:MF) hierarchy: ‘Hormone binding’ (GO:0042562; \( P_{\text{Adj}} = 0.012 \)), ‘Inorganic anion transmembrane transporter activity’ (GO:0015103; \( P_{\text{Adj}} = 0.012 \)), ‘DNA-binding transcription activator activity’ (GO:0001216; \( P_{\text{Adj}} = 0.035 \)), ‘signalling receptor activator activity’ (GO:0030546; \( P_{\text{Adj}} = 0.039 \)) and ‘G protein-coupled receptor activity’ (GO:0004930; \( P_{\text{Adj}} = 0.039 \)) were among the most significantly enriched terms (Figure 7A; see also Supporting information, Table S5). Over-representation analysis identified 23 enriched terms in the GO:Biological Process (GO:BP) hierarchy. Terms related to ‘cellular response to hormone stimulus’ (GO:0010222; \( P_{\text{Adj}} = 3.89 \times 10^{-3} \)), ‘inorganic ion transmembrane transporter activity’ (GO:0015103; \( P_{\text{Adj}} = 4.16 \times 10^{-3} \)), ‘signal transduction’ (GO:0007218; \( P_{\text{Adj}} = 7.51 \times 10^{-3} \)) and ‘regulation of response to hormone stimulus’ (GO:0051100; \( P_{\text{Adj}} = 1.67 \times 10^{-2} \)) were among the most significantly enriched terms (Figure 7B; see also Supporting information, Table S5).

Moreover, genes related to polymerase-II related transcription activity (GO:0007175; \( P_{\text{Adj}} = 0.012 \)), ‘proteolysis’ (GO:0007175; \( P_{\text{Adj}} = 0.012 \)) and ‘DNA-binding’ (GO:0000080; \( P_{\text{Adj}} = 0.012 \)) were among the most significantly enriched terms (Figure 7C; see also Supporting information, Table S5).
**FIGURE 3** Top expressed genes in the SON from control (euhydrated) rats according to their actions as transcription factors or their pharmacological classifications. A summary is provided of the top 5% most highly expressed genes. Data are presented as normalised counts ± SD. GPCRs, G protein coupled receptors. Classifications examined were transcription factors (A), enzymes (B), endogenous peptides (C), GPCRs (D), transporters (E), catalytic receptors (F), ion channels (G) and other targets (H).

Pcsk1 \((\log_{2}FC = 0.75; P_{\text{adj}} = 2.48 \times 10^{-18})\) and Pam \((\log_{2}FC = 0.63; P_{\text{adj}} = 5.09 \times 10^{-39})\), also known to be involved in neuropeptide processing. The top 5 most significant DEGs associated with neuropeptide signalling were Ecel1 \((\log_{2}FC = 0.65; P_{\text{adj}} = 1.97 \times 10^{-12})\), Mchr1 \((\log_{2}FC = 0.61; P_{\text{adj}} = 2.04 \times 10^{-10})\), Pdyn \((\log_{2}FC = 1.03; P_{\text{adj}} = 1.16 \times 10^{-22})\), Rxfp3 \((\log_{2}FC = 1.24; P_{\text{adj}} = 7.33 \times 10^{-28})\) and Glp1r \((\log_{2}FC = 0.96; P_{\text{adj}} = 1.33 \times 10^{-21})\) transcripts. In addition, GO:BP showed enrichment related to cAMP mediated signalling \(\langle\text{cAMP-mediated signalling}\rangle\), GO:0019933, \(P_{\text{adj}} = 1.40 \times 10^{-2}\) ‘response to cAMP’, GO:0051591, \(P_{\text{adj}} = 1.88 \times 10^{-3}\) and ‘ERK1 and ERK2 cascade’ (GO:0070371, \(P_{\text{adj}} = 3.63 \times 10^{-2}\) were also among the enriched terms in GO:BP (Figure 7B; see also Supporting information, Table S5). Notably, Fosl1 \((\text{LFC} = 3.40; P_{\text{adj}} = 1.25 \times 10^{-64})\) was associated with most of the enriched GO:BP terms and was the
most up-regulated gene in the dehydrated SON (Figure 7B). Note that it has previously been proposed that cAMP pathways are important in the SON response to dehydration.\(^2\)

Over-representation analysis against the KEGG and Reactome databases (Figure 7C,D; see also Supporting information, Table S5) identified Neuroactive ligand-receptor interaction (rno04080, \(P_{\text{Adj}} = 2.06 \times 10^{-4}\)) and IL-17 signalling pathway (rno04657, \(P_{\text{Adj}} = 4.37 \times 10^{-2}\)) as enriched KEGG pathways, whereas GPCR ligand binding (R-RNO-500792, \(P_{\text{Adj}} = 8.68 \times 10^{-3}\)) was the only enriched Reactome pathway. This highlighted a number of potentially interesting GPCR genes enriched as a consequence dehydration: Grpr, Nmnr2, Mchr1, Glp1r, Cckbr, Ptgr, Ptger4, Mc4r, Rxfp3, Scrt, Chrnna5, Gabrd, Gabq, Chrnrb4, F2r, Gpr68 and Opn3. The robust up-regulation of a number of receptors for gut peptides is intriguing and suggests that the SON might be a novel brain region involved in food intake and appetite.

To further explore the molecular pathways undergoing change in dehydration we performed GSEA\(^{31}\) that screens for functional enrichment of molecular signatures using the entirety of expressed genes in the dehydrated SON. The rationale behind GSEA analysis is that a lot of small expression changes (not necessarily breaching the significance threshold to be flagged as differentially expressed) of a set of genes may add up to large net effect on a specific molecular function.

Gene-set enrichment analysis identified three GO:MF terms including ‘peptide binding’ (GO:0042277, normalised enrichment score \([\text{NES}] = 1.708, P_{\text{Adj}} = 2.03 \times 10^{-5}\) ), ‘peptide hormone binding’ (GO:0017046, \(P_{\text{Adj}} = 4.37 \times 10^{-2}\)) and ‘hormone binding’ (GO:0017046, \(P_{\text{Adj}} = 3.32 \times 10^{-2}\)) that showed strong positive enrichment (up-regulation) in dehydration, whereas ‘Structural constituent of ribosome’ (GO:0003735, \(P_{\text{Adj}} = 8.24 \times 10^{-3}\)) was the most significant enriched term indicating
FIGURE 5 Top supraoptic nucleus (SON) differentially expressed genes (DEGs) following dehydration. A, The top 100 DEGs in the SON as a consequence of dehydration. The top 100 genes were selected by listing all identified DEGs by $P_{\text{Adj}}$ listed ($P_{\text{Adj}}$) value in increasing order. Notably, 97 out of 100 most-significant DEGs are up-regulated in dehydration. DEGs are ordered by expression change (log 2 fold-change [FC]) in dehydration in decreasing order. Colour scaling of the bars represents $P_{\text{Adj}}$ values of the DEGs. Colour scaling of the dots represents log2 fold-change expression values. Dots are sized according to transcript abundance (expression) measured by normalised read counts aligned to each gene. B, Top 100 DEGs in the SON following dehydration (relative change in dehydration). The Top 100 genes were selected by filtering all DEGs ($P_{\text{Adj}} < 0.05$) and selecting the Top 100 genes with the greatest log2 fold-change values. Genes are listed and coloured by log2 fold-change in decreasing order. Colours of the bars represent $P_{\text{Adj}}$ values. Size of the dots represent average expression of the transcript between euhydrated and dehydrated conditions measured by normalised uniquely-mapped reads. C, Top 100 DEGs in SON in dehydration (absolute change in dehydration). We observed that log2 fold-change (relative expression in dehydration compared to the control) favours lowly expressed transcripts over the abundant genes (B) thus we selected the top 100 genes with the greatest expression change favouring highly expressed genes. For this, genes were filtered for differential expression ($P_{\text{Adj}} < 0.05$) and the difference between the average normalised read count between the control and dehydrated conditions (delta) was calculated. The top 100 genes with the greatest delta values were selected. This represents an expression change in dehydration measured by absolute units (uniquely mapped reads) rather than relative change from the control condition. This analysis highlights the greatest expression changes in dehydration of highly expressed transcripts.

A strong negative enrichment (down-regulation) of associated genes (Figure 8A).

Gene-set enrichment analysis using the GO:BP hierarchy identified 44 functionally enriched terms. Notably, significantly enriched terms included ‘response to steroid hormone’ (GO:0048545; NES = 1.80, $P_{\text{Adj}} = 1.66 \times 10^{-5}$), ‘response to peptide hormone’ (GO:0043434, NES = 1.79, $P_{\text{Adj}} = 1.81 \times 10^{-5}$), ‘response to cAMP’ (GO:0051591, NES = 1.98, $P_{\text{Adj}} = 3.36 \times 10^{-3}$), ‘response to endoplasmic reticulum stress’ (GO:0034976, NES = 1.64, $P_{\text{Adj}} = 3.87 \times 10^{-2}$) and ‘regulation of transcription from RNA polymerase II promoter in response to stress’ (GO:0043618, NES = 1.92, $P_{\text{Adj}} = 3.99 \times 10^{-3}$) (Figure 8B).

Lastly, we performed GSEA against the KEGG and Reactome pathway databases. The most significantly enriched KEGG pathway was ‘Ribosome’ (rno03010, NES = −1.97, $P_{\text{Adj}} = 6.44 \times 10^{-5}$) and indicated a global down-regulation of translation related genes (Figure 8C). Other enriched KEGG pathways included ‘Relaxin signaling pathway’ (rno04926, NES = 2.01, $P_{\text{Adj}} = 1.49 \times 10^{-3}$), ‘PI3K-Akt signaling pathway’ (rno04151, NES = 1.65, $P_{\text{Adj}} = 2.07 \times 10^{-2}$), ‘Neuroactive ligand-receptor interaction’ (rno04080, NES = 1.63, $P_{\text{Adj}} = 2.07 \times 10^{-2}$) and ‘Protein processing in endoplasmic reticulum’ (rno04141, NES = 1.72, $P_{\text{Adj}} = 2.07 \times 10^{-2}$). Notably, with the exception of ‘Ribosome’ pathway, all enriched KEGG pathways indicated positive enrichment (up-regulation) of associated set of genes (Figure 8C).

No Reactome pathways passed multiple comparison corrected significance threshold (Benjamini-Hochberg; $P_{\text{Adj}} < 0.05$) for gene set enrichment. However, 25 Reactome pathways indicated significant enrichment of the pathway prior to multiple comparison correction including ‘Signaling by GPCR’ (R-RNO-372790, NES = 1.52, $P = 0.0012, P_{\text{Adj}} = 0.13$), ‘G alpha(s) signalling events’ (R-RNO-372790, NES = 1.71, $P = 0.0015, P_{\text{Adj}} = 0.13$), ‘Post-translational protein phosphorylation’ (R-RNO-8957275, NES = 1.73, $P = 0.0016, P_{\text{Adj}} = 0.13$) and ‘Eukaryotic Translation Initiation’ (R-RNO-72613, NES = −1.51, $P = 0.0029, P_{\text{Adj}} = 0.13$) (see Supporting information, Table S5).
Together, GSEA indicated pathways related to peptide and steroid hormone signalling, activation of transcription, and cAMP, IL-17, PI3K-Akt and relaxin intracellular signalling were positively enriched (ie activated at the steady-state RNA level) in response to water deprivation. Interestingly, GSEA revealed a core set of genes encoding ribosomal proteins related to
FIGURE 6 Gene expression changes in the rat supraoptic nucleus (SON) as a consequence of dehydration categorised according to functions as transcription factors or pharmacological classification. The table summarises the total number of statistically significant regulated genes (P_adj < 0.05), the number of genes with fold change (FC) bigger than 25% (P_adj < 0.05; FC > 25%) and the number of genes with fold change bigger than 50% (P_adj < 0.05; FC > 50%). The graphs present the 367 genes with a fold change greater than 25%. Data are presented as the power of 2 of the mean of normalised counts ± SD. GPCRs, G protein coupled receptors; WD, water deprived. Classifications examined were transcription factors (A), enzymes (B), endogenous peptides (C), GPCRs (D), transporters (E), catalytic receptors (F), ion channels (G) and other targets (H).

FIGURE 7 Pathway analysis of supraoptic nucleus (SON) genes differentially regulated by dehydration. A, Dot plot of over-represented Gene Ontology (GO):Molecular Function (MF) terms and their associated differentially expressed genes (DEGs). Over-represented GO:MF terms are listed on the y-axis by P_adj value from top to bottom in increasing order. The top 5 most-significant associated DEGs (by P_adj) of each over-represented GO term are shown as dots on the x-axis coloured by the direction of expression change in dehydration (log2 fold-change [FC]) and sized according to transcript abundance (expression) measured by normalised read counts aligned to each gene. B, The top 23 over-represented GO:Biological Process (BP) terms. A truncated list of 131 over-represented GO:BP terms is shown. A complete list is provided in the Supporting information (Table S5). C, Over-represented Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. DEGs (P_adj < 0.05) associated with each term are displayed on the x-axis. D, Over-represented Reactome pathways. DEGs (P_adj < 0.05) associated with each term are displayed on the x-axis.

translation being supressed in response to dehydration (enriched pathways: GO:0003735, rno03010, R-RNO-72613). This is particularly surprising because the majority of enriched pathways identified by GSEA indicated strong positive enrichment in line with SON activation by dehydration. We suggest that this could be explained as a secondary response to the 72 hours of dehydration protocol in which transcription of translation machinery is supressed after a prolonged period of activation. Complete lists of pathway analysis results can be found in the Supporting information (Table S5).
3.5 | Transcriptional regulatory networks enriched in the SON in response to dehydration

Regulons consists of a transcription factor (TF) and its effector genes forming a transcriptional regulatory network. The TF representing the regulon can either have a positive or a negative interaction with the target genes comprising the regulon, acting via direct (TF-target) or indirect (TF-TF-target) regulation (Figure 9A). Thus, a subtle change in regulon TF expression can have a large net effect on a specific function by affecting multiple downstream mediators. We have performed transcriptional regulatory network analysis using Wistar SON transcriptome data to identify transcriptional...
regulatory networks affected by dehydration. This analysis identified Tcf25, Sin3a, Hes6, Rxrb, Rbck1, Zfp467, Zfp955a and Tcfp2 as transcription factors constituting regulatory networks in the Wistar SON (Figure 10B; see also Supporting information, Table S6). To assess whether any of the identified Regulons undergo changes in response to 72 hours of dehydration, we performed Master Regulator Analysis (MRA) and two-tailed GSEA. We found eight regulons affected by dehydration by MRA (Sin3a, \( P_{\text{Adj}} = 5.60 \times 10^{-4} \); Rxrb, \( P_{\text{Adj}} = 1.40 \times 10^{-17} \); Hes6, \( P_{\text{Adj}} = 3.60 \times 10^{-3} \); Tcfp2, \( P_{\text{Adj}} = 2.80 \times 10^{-3} \); Tcf25, \( P_{\text{Adj}} = 6.60 \times 10^{-5} \); and Zfp955a, \( P_{\text{Adj}} = 3.90 \times 10^{-2} \)) (Figure 10B), whereas five regulons passed multiple comparison corrected significance threshold using GSEA (Sin3a, \( \text{ES} = -0.66, P_{\text{Adj}} = 0.019 \); Rxrb, \( \text{ES} = -0.59, P_{\text{Adj}} = 0.014 \); Tcf25, \( \text{ES} = -0.59, P_{\text{Adj}} = 0.019 \); Zfp955a, \( \text{ES} = -0.7, P_{\text{Adj}} = 0.019 \); and Rbck1, \( \text{ES} = -0.71, P_{\text{Adj}} = 0.034 \)) (Figure 9B,C). Four of eight identified TFs were significantly down-regulated in response to dehydration (Hes6, Rxrb, Rbck1 and Zfp467), whereas all indicated a trend, albeit small, towards down-regulation in response to water deprivation. Notably, Tcf25 is a highly expressed transcript (similar to Rpl4, Rpl3 or Myl6) in Wistar SON, whereas other regulon TFs had intermediated expression values (see Supporting information, Table S3).

To visualise the change in gene expression of the enriched regulons in dehydration, we plotted the log2FC values of genes comprising the regulatory networks for each regulon (Figure 9D). All identified regulons had a fairly small size of their regulatory networks with only positive interactions to the effector genes. All genes regulated by the regulon TFs indicated down-regulation in response to...
water deprivation. For most genes, the observed expression change was small and not always pass the multiple comparison corrected significance threshold. Notably, the Tcf25 regulon was the largest with 42 genes in its network, whereas the Tcfp2 and Zfp955a regulons were much smaller and consisted of only 12 genes.

To gain insight into the molecular functions of these regulons, we performed pathway and GO analysis using all unique genes associated with all TFs, and with each individual transcriptional regulatory network. Examining all genes in the network, over-representation analysis identified ‘structural constituent of ribosome’ (GO:MF; GO:0003735; \( P_{\text{Adj}} = 0.059\)), ‘Ribosome’ (KEGG; rno03010; \( P_{\text{Adj}} = 0.016\)), ‘Mitophagy – animal’ (KEGG; rno04137; \( P_{\text{Adj}} = 0.031\)) and pathways associated with ‘Eukaryotic Translation Initiation’ (Reactome; R-RNO-72613; \( P_{\text{Adj}} = 5.79 \times 10^{-5}\)) and ‘Translation’ (Reactome; R-RNO-72766; \( P_{\text{Adj}} = 5.93 \times 10^{-5}\)) to be enriched among the regulon genes (see Supporting information, Table S8). Focussing on each individual TF, we found that the Tcf25, Hes6, Sin3a and Rxrb regulons were similarly enriched in terms related to eukaryotic translation, whereas the Rbck1 network showed enrichment related to ‘histone methyltransferase activity’ (GO:MF; GO:0042054; \( P_{\text{Adj}} = 0.018\)) (see Supporting information, Table S7).

We note that the largest regulon, Tcf25, contained two other TFs (Hes6 and Sin3a) represented in its regulatory network. Similarly, the Hes6 and Sin3a regulons contained Tcf25 within their gene networks implying a reciprocal regulatory mechanism in which down-regulation of one transcription factor would suppress its own expression in a negative feedback loop. The aforementioned regulons shared a considerable overlap in the number of effector genes in their regulatory networks (Figure 9E; see also Supporting information, Table S7). Interestingly, the pathway analysis on the Tcf25, Hes6 and Sin3a regulatory networks revealed all three TFs to be associated with eukaryotic translation (see Supporting information, Table S7). This coincides with the pathway analysis results that ‘structural constituent of ribosome’ and ‘ribosome’ pathways indicating strong negative enrichment (suppression) by GSEA in the context of global phenotypic switch in dehydration (Figure 8A,C). Both these findings contrast with the majority of DEGs increasing their expression in dehydration and the majority of molecular pathways identified by GSEA indicating strong positive enrichment (Figure 8). We suggest that the described regulons are at least in part responsible for the observed down-regulation of translation associated genes in the Wistar SON in response to water deprivation.

This down-regulation in the expression of a large number of genes involved in ribosome function and translation is counter-intuitive given the increased demand for the production of secreted peptide hormones following dehydration. There are a number of possibilities. First, translation may be suppressed to avoid endoplasmic reticulum (ER) overload and hence the triggering of ER stress. Second, down-regulation of these genes may not necessarily suppress translation and may in fact enhance it. Third, it is possible that this change in the balance of translation factors favours the translation of certain transcripts, at the same time as perhaps reducing the translation of others.

### 3.6 | Common dehydration genes

To reveal the core set of gene transcripts differentially expressed in the SON as a consequence of dehydration, we compared the newly generated Wistar RNA-seq dataset with the previously published\(^{17}\) and publicly available (NCBI GSE3110) SON microarray data in dehydrated condition in Sprague-Dawley rats. It is important to note that the two experiments differ in rat strain (Sprague-Dawley vs Wistar), tissue sampling technique (Sprague-Dawley – crude hypothalamic area dissection; Wistar – frozen-section punching) and gene expression detection technology (Sprague-Dawley – microarray; Wistar – RNA-seq). The bioinformatics pathway used to enable this comparison is illustrated in the Supporting information (Figure S10).

Overall, fewer DEGs were identified in Sprague-Dawley (\( n = 1244; P_{\text{Adj}} < 0.005\)) compared to Wistar (\( n = 2259; P_{\text{Adj}} < 0.05\)) rats (Figure 10A). To further compare the Wistar and SD datasets, we plotted DEGs identified in each experiment (\( P_{\text{Adj}} < 0.05\)) according to their expression change (log2FC) (Figure 10B; see also Supporting information, Figure S11). The RNA-seq approach detected more small expression changes compared to microarray as the greatest difference between the datasets was in genes with expression change values of \( 0.5 < \text{log2FC} < -0.5\) (Figure 10B). In addition, the Wistar (RNA-seq) dataset displayed a left skewed \( P\) adjusted values distribution that was expected given a major transcriptional change that occur in SON in response to dehydration. This showed that a large number of alternative hypotheses (DEGs) were confidently identified after false positive removal by multiple comparison correction (see Supporting information, Figure S11B). By contrast, the Sprague-Dawley (microarray) dataset displayed uniform \( P\) adjusted values distribution suggesting that a large number of true positive hits were not identified and possibly removed by multiple comparison correction as a result of the poor performance of initial differential expression calling (assuming the same number of genes respond to dehydration in both strains).

Next, we analysed the overlap between the Wistar and Sprague-Dawley datasets and found 504 genes to commonly change their expression in response to dehydration (Figure 9A). We suggest these genes to represent a core set of genes that change their expression in the rat hypothalamic SON in response to water deprivation. The majority of common DEGs had low expression change values (\( 0.5 < \text{log2FC} < -0.5\)) (Figure 10B,D) and a high confidence of DEG calling (\( P_{\text{Adj}} < 0.05\)) (Figure 10C; see also Supporting information, Figure S11C). Interestingly, a much greater number of common DEGs were up-regulated in dehydration (422 up-regulated vs 82 down-regulated genes) (Figure 10B,D). Common DEGs showed a high degree of correlation in their expression change in response to dehydration (Spearman rho = 0.78; \( P = 2.2 \times 10^{-16}\)) (Figure 10C). Notably, several of common DEGs had an opposite direction of
expression change (Figure 10C). Nine transcripts (Fam180a, Svil, Slc6a20, Mrvi1, Plscr1, Cavin2, Gramd2b, Finb and Des) were detected to be down-regulated in SD rats, whereas it was up-regulated in the Wistar rats. Only two transcripts (Wdr6, Rbfox1) were detected to have an opposite mismatch of expression change. Lastly, we used a volcano plot to visualise common DEGs in dehydration (Figure 10D). These genes include Fosl1, Tnfaip6, Dnase2b, Vgf, Arhgdib, Caprin2, Creb3I1, Opn3, Slc12a1, Dyps, Plaur, Ptges, Procr and Igfbp5. Lastly, we outlined how many of the common DEGs were among the genes with the greatest expression change in response to dehydration as shown in Figure 5 (see also Supporting information, Figure S11D,E). A full list of DEGs common and unique to each analysis can be found in the Supporting information (Table S8). Many of these genes have been previously validated using either a quantitative reverse transcription polymerase chain reaction, or by in situ hybridisation, in Sprague-Dawley and Wistar rats (Table 1).

4 | CONCLUSIONS AND PERSPECTIVES

We have comprehensively catalogued the polyadenylated transcriptome of the euhydrated male Wistar Han rat SON, and we have described how this transcriptome changes following 72 hours of dehydration. Deep mining of our data has revealed a plethora of genes and gene categories that are hence implicated in the function related plasticity exhibited by the physiologically challenged SON.

We hope that these datasets will open new avenues of HNS research. The genome sequencing projects revealed that approximately 22 000 protein coding genes are needed to make a mammal but a mere 5% of genes dominate 70% of neuroscience publications. The remaining 95% of genes have been defined as the ignorome. Stoeger et al investigated the reasons why potentially important genes are ignored. They found that "biological research is primarily guided by a handful of generic chemical and biological characteristics of genes, which facilitated experimentation during the 1980s and 1990s, rather than the physiological importance of individual genes." In our view, this convergence is detrimental to science and hinders discovery. This work was intended to serve as a resource for the diverse neuroendocrine research community, persuading it to acknowledge the unexplored frontiers of the HNS. We hope that the revelation of a core set of genes and molecular pathways governing the osmotic response, many with no characterised roles in the SON, will encourage the neuroendocrine community to pursue new investigations into the ‘known-unknowns’ presented in this study.

5 | STUDY LIMITATIONS

5.1 | RNA-seq

We sequenced the polyadenylated transcriptomes of the SON, thus focussing, in the main, on transcripts (mRNAs) that code for proteins. Most non-coding RNAs will have been excluded from our analysis.

5.2 | The dehydration model

Three days of complete fluid deprivation has traditionally been the standard hyperosmotic stimulus used to elicit robust activation of AVP MCNs. Although effective and supported by a huge literature, this model has been criticised as being stressful to the animal and insufficiently specific because many different physiological parameters (including circulating blood volume and changed energy metabolism) and molecular pathways are simultaneously activated, hindering precise interpretation.

| Symbol | Sprague-Dawley | Wistar Han |
|--------|----------------|------------|
| Atf4   | 18, 21, 24, 31 |
| Atf5   | 21            |
| Ap1s2  | 31            |
| Arhgdib| 35            |
| Atp1a2 | 31            |
| Azin1  | 19, 20, 28    |
| BIP (Hspa5, Grp78) | 24 |
| Caprin2| 30, 31        |
| Cebp   | 18, 26        |
| Chop   | 24            |
| Creb3I1| 18, 23, 24    |
| Fos    | 61            |
| Fosl1  | 26            |
| Giot1  | 18, 20, 21    |
| Giot2b | 18            |
| Hbb    | 31            |
| Igfbp2 | 31            |
| Itgb1  | 21            |
| Nab1   | 21            |
| Nampt  | 19            |
| Nr4a1  | 26            |
| Opn3   | 21, 31        |
| Pcp4   | 31            |
| Pcsk1  | 27            |
| Pdyn   | 31            |
| Procr  | 21            |
| Psph   | 21            |
| Ptges  | 19            |
| Ran    | 31            |
| Rasd1  | 32            |
| Rgs5   | 25            |
| Slc12a1| 29            |
| Tnfaip6| 19, 20        |
| Trpv2  | 21            |
| Vgf    | 21            |

TABLE 1 | Published validation by independent methods of Sprague-Dawley and Wistar supraoptic nucleus transcriptome data
In the present study, we have focussed on the transition from euhydration to chronic dehydration. It would be of great interest to determine how the transcriptome changes when animals are allowed to recover during a period of rehydration.

5.3 | Cellular heterogeneity

Although every care was taken to accurately punch the SON from frozen sections, it is inevitable that some surrounding tissue will be included. This will increase the noise of the analysis and may introduce signal that is not pertinent to the SON. Furthermore, the bulk SON sequencing carried includes transcriptome information from every cell type in this nucleus, not only the oxytocinergic and vasopressinergic MCNs, but also the surrounding glia, microglia, some interneurones, and vessels and the blood therein. We await with great interest the inevitable single cell RNA-seq analysis of the euhydrated and dehydrated SON, which will be highly informative regarding the transcriptomic responses of these different cell types. Furthermore, it is to be expected that the MCNs themselves will exhibit an intrinsic diversity with respect to basal gene expression patterns and responses to physiological cues that will have functional implications. We note that, in addition to secretion from the PP terminals into the systemic circulation, AVP and OXT can be also released within the hypothalamic SON and PVN from the dendrites and cell bodies of the MCNs, hence acting as neurotransmitters. These mechanisms are usually independent of the axonal release of AVP and OXT into the blood stream and dendritic release of these peptides can be achieved by intracellular Ca²⁺ mobilisation without electrical firing of the neurones.⁶⁷,⁶⁸ Somatodendritic release of AVP and OXT has important autocrine effects in MCNs and paracrine effects in nearby cells regulating a variety of functions such as the blood pressure⁶⁹ or the milk-ejection reflex.⁷⁰ In addition, MCNs from the SON and PVN also send extra-neurohypophysial collateral projections to other brain areas contributing to central actions and affecting complex outputs, such as regulating pain perception,⁷¹ affective/emotional states,⁷²,⁷³ social behaviour,⁷⁴,⁷⁵ fear,⁷⁶,⁷⁷ and general anaesthesia and sleep.⁷⁸ It will be of great interest to identify individual MCN transcriptome signatures that mediate to these diverse physiological and behavioural roles.

5.4 | GO analysis

GO and pathway analysis is limited by the fact that most genes are not annotated with functional GO and pathway information. The extent of the ignorome, by definition, impacts upon the biological value of GO analyses. It has been shown that there continues to be a strong annotation bias in the GO annotations where 58% of the annotations are for 16% of the human genes.⁷⁹ Moreover, genes have independent GO, KEGG and Reactome annotations, making it difficult to compare the enrichment results linking different databases. GO, as the name implies, is an unordered and unstructured set of genes associated with a specific biological process or mechanism. Although inherently sensitive at detecting any type of enrichment, gene-set (ontology) analysis results are sometimes difficult to be put into biological context. By contrast, the KEGG and Reactome databases contain information on the interaction of genes, gene products or metabolites associated with a specific mechanisms or pathways. Structured into pre-defined and directional pathways, and based on prior known gene interactions, pathway analysis may struggle at identifying enrichment if the genes affected by the treatment are not previously known to interact.

Enrichment analysis is also dependent on the probability based (ORA) and functional class scoring (GSEA) classes of methods.⁸⁰ Although sensitive and widely used, ORA is biased and depends heavily on the criteria used to select the differential expression threshold. In addition, ORA-based approaches cannot differentiate whether input genes profoundly change expression or just barely breach the significance threshold. To account for this, GSEA can be employed to incorporate information about all genes expressed in a given sample. GSEA accounts for the level of differential expression by weighting the expression change to calculate the enrichment score (heavily up-regulated/down-regulated genes contribute more to the final enrichment score than the genes that show only marginal change in expression). The rationale behind this approach is that cumulative, albeit individually small, changes in gene expression of multiple functionally related genes add up to a bigger, physiologically significant effect on downstream function of these genes. Although sensitive and less biased at detecting functional enrichment, in the context of downstream analysis, the implications of minute expression change of multiple genes are difficult to interpret and challenging to test experimentally.

5.5 | Regulon analysis

Prediction of transcriptional regulatory networks requires a defined list of transcription factors that are screened against the gene-expression profile of a given sample. In our analysis, we identified expressed transcription factors using GO annotations. However, it should be noted that the identification of transcription factors is based on binding site predictions and not always followed by functional characterisation. Furthermore, regulons analysis was carried out on a limited number of samples used in the experiment (n = 5). Although the sample size was suitable for the differential expression analysis, experimental power of the regulon analysis could be further improved by a larger sample size. Lastly, all identified transcription factors representing a transcriptional network had only positive interactions with their regulatory genes and were relatively small in size. This represents a limitation for the two-way GSEA analysis that was able to assess only unidirectional enrichment associated with the TF expression.

ACKNOWLEDGEMENTS

This work was supported by the Biotechnology and Biological Sciences Research Council (BB/J005452/1, DM, CCTH; BB/J015415/1, DM, MPG, BB/R016879/1, DM), the Medical Research
CONFICT OF INTERESTS
The authors declare that they have no conflicts of interest.

AUTHOR CONTRIBUTIONS
Audrys Pauža: Data curation; Formal analysis; Software; Writing – review & editing. André S Mecawi: Formal analysis; Writing – original draft; Writing – review & editing. Alex Paterson: Data curation; Formal analysis; Software; Writing – review & editing. Charles Hindmarsh: Formal analysis; Writing – review & editing. Mingkwon P Greenwood: Conceptualisation; Formal analysis; Investigation; Methodology; Project administration; Supervision; Writing – review & editing. David Murphy: Conceptualisation; Funding acquisition; Project administration; Supervision; Writing – original draft; Writing – review & editing. Michael P Greenwood: Conceptualisation; Formal analysis; Investigation; Methodology; Project administration; Supervision; Writing – review & editing.

PEER REVIEW
The peer review history for this article is available at https://pub&&!ns.com/publon/10.111/j.nje.13007.

DATA AVAILABILITY STATEMENT
In addition to the processed data reported in the manuscript, all raw data are available from the corresponding author upon reasonable request.

ORCID
Audrys G. Pauža https://orcid.org/0000-0001-9635-8872
André Souza Mecawi https://orcid.org/0000-0003-4517-6221
David Murphy https://orcid.org/0000-0003-2946-0353

REFERENCES
1. Mecawi AE, Ruginsk SG, Elias LL, Varanda WA, Antunes-Rodrigues J. Neuroendocrine regulation of hydromineral homeostasis. J Comp Physiol. 2015;5:1465-1516.
2. Murphy D, Antunes-Rodrigues J, Gainer H. Osmoregulation. In: Murphy D, Gainer H, eds. Molecular Neuroendocrinology: From Genome to Physiology. Hoboken, NJ: Wiley-Blackwell; 2016:331-354.
3. Bargmann W. Neurosecretion. Int Rev Cytol. 1966;19:183-201.
4. Brownstein MJ, Russell JT, Gainer H. Synthesis, transport, and release of posterior pituitary hormones. Science. 1980;207:373-378.
5. Burbach JP, Luckman SM, Murphy D, Gainer H. Gene regulation in the magnocellular hypothalamo-neurohypophysial system. Physiol Rev. 2001;81:1197-1267.
6. Bourque CW. Central mechanisms of osmosensation and systemic osmoregulation. Nat Rev Neurosci. 2008;9:519-530.
7. McKinley MJ, Mathai ML, McAllen RM, et al. Vasopressin secretion: osmotic and hormonal regulation by the lamina terminals. J Neuroendocrinol. 2004;16:340-347.
8. Breyer MD, Ando Y. Hormonal signaling and regulation of salt and water transport in the collecting duct. Annu Rev Physiol. 1994;56:719-739.
9. Ares GR, Caceres PS, Ortiz PA. Molecular regulation of NKCC2 in the thick ascending limb. Am J Physiol. 2011;301:F1143-F1159.
10. Murphy D, Carter D. Vasopressin gene expression in the rodent hypothalamus: transcriptional and post-transcriptional responses to physiological stimulation. Mol Endocrinol. 1990;4:1051-1059.
11. Kondo N, Arima H, Banno R, Kuwahara S, Sato I, Oiso Y. Osmoregulation of vasopressin release and gene transcription under acute and chronic hypovolemia in rats. Am J Physiol Endocrinol Metab. 2004;286:E337-E346.
12. Sherman TG, Civelli O, Douglass J, Herbert E, Watson SJ. Coordinate expression of hypothalamic pro-dynorphin and pro-vasopressin mRNAs with osmotic stimulation. Neuroendocrinology. 1986;44:222-228.
13. Huang W, Lee SL, Arnason SS, Sjoquist M. Dehydration natriuresis in male rats is mediated by oxytocin. Am J Physiol. 1996;270:R427-R433.
14. Hatton GI. Function-related plasticity in hypothalamus. Ann Rev Neurosci. 1997;20:375-379.
15. Theodosis DT, El Majdoubi M, Pierre K, Poulin DA. Factors governing activity-dependent structural plasticity of the hypothalamo-neurohypophysial system. Cell Mol Neurobiol. 1998;18:285-298.
16. Sharman G, Ghorbel M, Leroux M, Beaucourt S, Wong LF, Murphy D. Deciphering the mechanisms of homeostatic plasticity in the hypothalmo-neurohypophysyal system–genomic and gene transfer strategies. Prog Brain Res Mol Biol. 2004;84:151-182.
17. Hindmarsh C, Yao S, Beighton G, Paton J, Murphy D. A comprehensive description of the transcriptome of the hypothalamoneurohypophysyal system in euhydric and dehydrated rats. Proc Natl Acad Sci USA. 2006;103:1609-1614.
18. Qiu J, Yao S, Hindmarsh C, Antunes V, Paton J, Murphy D. Transcription factor expression in the hypothalmo-neurohypophysyal system of the dehydrated rat; up-regulation of Gonadotrophin inducible transcription factor 1 mRNA is mediated by cAMP-dependent protein kinase A. J Neurosci. 2007;27:2196-2203.
19. Stewart L, Hindmarsh CCT, Qiu J, Tung Y-CL, Yeo GSH, Murphy D. Hypothalamic transcriptome plasticity in two rodent species reveals divergent differential gene expression but conserved pathways. J Neuroendocrinol. 2011;23:177-185.
20. Qiu J, Hindmarsh C, Yao S, Tasker J, Murphy D. Transcriptomic analysis of the osmotic and reproductive remodelling of the female rat supraoptic nucleus. Endocrinology. 2011;152:3483-3491.
21. Greenwood M, Mecawi AS, Hoe SZ, et al. A comparison of physiological and transcriptome responses to water deprivation and salt loading in the rat supraoptic nucleus. Am J Physiol. 2015;308:R559-R568.
22. Yue C, Mutsuga N, Verbalis J, Gainer H. Microarray analysis of gene expression in the supraoptic nucleus of normoosmotic and hypoosmotic rats. Cell Mol Neurobiol. 2006;26:959-978.
23. Greenwood M, Bordieri L, Greenwood MP, et al. Transcription factor CREB3L1 regulates vasopressin gene expression in the rat hypothalamus. J Neurosci. 2014;34:3810-3820.
24. Greenwood M, Greenwood MP, Paton JFR, Murphy D. Transcription factor CREB3L1 regulates endoplasmic reticulum stress response genes in the osmotically challenged rat hypothalamus. PLoS One. 2015;10:e0124956.
25. Greenwood M, Greenwood MP, Mecawi AS, et al. Transcription factor CREB3L1 mediates cAMP and glucocorticoid regulation of arginine vasopressin gene transcription in the rat hypothalamus. Mol Brain. 2015;8:68.
26. Greenwood MP, Greenwood M, Gillard BT, Devi CR, Murphy D. Regulation of CAMP responsive element binding protein 3-like 1 (Creb3l1) expression by orphan nuclear receptor Nr4a1. Front Mol Neurosci. 2018;10:413.

27. Greenwood M, Paterson A, Rahman PA, et al. Transcription factor Creb3l1 regulates the synthesis of prohormone convertase enzyme PC1 in endocrine cells. J Neuroendocrinol. 2020;32:e12851.

28. Greenwood MP, Greenwood M, Paton J, Murphy D. Control of polyamine biosynthesis by anti-tumor inhibitor 1 is important for transcriptional regulation of arginine vasopressin in the male rat hypothalamus. Endocrinol. 2015;156:2905-2917.

29. Konopacka A, Qiu J, Yao ST, et al. Osmoregulation requires brain expression of the renal Na-K-2Cl co-transporter NKCC2. J Neurosci. 2015;35:5144-5155.

30. Konopacka A, Greenwood M, Loh S-Y, Paton J, Murphy D. RNA binding protein Caprin-2 is a pivotal regulator of the central osmotic defense response. eLife. 2015;4:e09656.

31. Loh S-Y, Jahans-Price T, Greenwood MP, et al. Unsupervised network analysis of the plastic supraoptic nucleus transcriptome predicts Caprin-2 regulatory interactions. eNeuro. 2017;4(6):ENEURO.0243-17.2017.

32. Greenwood MP, Greenwood M, Mecawi A, Antunes-Rodrigues J, Paton J, Murphy D. Rasd1, a small G protein with a big role in the hypothalamic response to neuronal activation. Mol Brain. 2016;9:1.

33. Ueta Y. AZ street in vasopressin synthesis. Endocrinology. 2015;156:2749-2750.

34. Husain MK, Fernando N, Shapiro M, Kagan A, Glick SM. Radiomunoassay of arginine vasopressin in human plasma. J Clin Endocrinol Metab. 1973;37:616-625.

35. Andrews S. (2010-2019) FastQC: a quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

36. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29:15-21.

37. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30:923-930.

38. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550.

39. Klang TF, Lau CY. Getting the most out of RNA-seq data analysis. PeerJ. 2015;3:e1360.

40. Yates AD, Achuthan P, Akanni W, et al. Ensembl 2020. Nucleic Acids Res. 2020;48:D682-D688.

41. Durincz S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nat Protoc. 2009;4:1184-1191.

42. Durincz S, Moreau Y, Kasprzyk A, et al. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. Bioinformatics. 2005;21:3439-3440.

43. Carlsson M. org.Rn.eg.db: Genome wide annotation for Rat. R package version 3.8.2. 2019.

44. Chen H, Boutros PC. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. BMC Bioinformatics. 2011;12:35.

45. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. Nat Genet. 2000;25:25-29.

46. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28:27-30.

47. Kanehisa M, Sato Y, Furumichi M, Morishima K, Tanabe M. New approach for understanding genome variations in KEGG. Nucleic Acids Res. 2019;47:D590-D595.

48. Yu G, Wang LG, Han Y, He QY. ClusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16:284-287.

49. Jassal B, Matthews L, Viteri G, et al. The reactome pathway knowledgebase. Nucleic Acids Res. 2020;48:D498-D503.

50. Yu G, He QY. ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. Mol Biosyst. 2016;12:477-479.

51. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005;102:15545-15550.

52. Fletcher MN, Castro MA, Wang X, et al. Master regulators of FGFR2 signalling and breast cancer risk. Nat Commun. 2013;4:2464.

53. Castro MA, de Santiago I, Campbell TM, et al. Regulators of genetic risk of breast cancer identified by integrative network analysis. Nat Genet. 2016;48:12-21.

54. Lambert SA, Jolma A, Campitelli LF, et al. The human transcription factors. Cell. 2018;175:598-599.

55. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Core Team; 2020.

56. Davis S, Meltzer PS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. Bioinformatics. 2007;14:1846-1847.

57. Ritchie ME, Phipson B, Wu D, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucl Acids Res. 2015;43:e47.

58. https://www.guidetopharmacology.org/

59. Armstrong JF, Faccenda E, Harding SD, et al. The IUPHAR/BPS Guide to PHARMACOLOGY in 2020: extending immunopharmacology content and introducing the IUPHAR/MMV Guide to MALARIA PHARMACOLOGY. Nucleic Acids Res. 2020;48:D1006-D1021.

60. Greenwood G, Romanova EV, Mecawi AS, et al. The effects of ageing on biosynthetic processes in the rat hypothalamic osmoregulatory neuroendocrine system. Neurobiol Aging. 2016;65:178-191.

61. Greenwood MP, Greenwood M, Gillard B, Loh SY, Paton JFR, Murphy D. Epigenetic control of the vasopressin promoter explains physiological ability to regulate vasopressin transcription in dehydration and salt loading states in the rat. J Neuroendocrinol. 2016;28(4):e12371. https://doi.org/10.1111/jne.12371.

62. Carro MS, Lim WK, Alvarez MJ, et al. The transcriptional network for mesenchymal transformation of brain tumours. Nature. 2010;463:318-325.

63. Stoeger T, Gerlach M, Morimoto RI, Nunes Amaral LA. Reply to “Far away from the lamppost”. PLoS Biol. 2018;16(12):e3000075.

64. Pandey AK, Lu L, Wang X, Hornayouni R, Williams RW. Functionally enigmatic genes: a case study of the brain ignoreme. PloS One. 2014;9:e88889.

65. Stoeger T, Morimoto RI, Nunes Amaral LA. Large-scale investigation of the reasons why potentially important genes are ignored. PloS Biol. 2018;16:e2006643.

66. Oprea TI, Jan L, Johnson GL, et al. Far away from the lamppost. PLoS Biol. 2018;16:e3000067.

67. Ludwig M, Sabatier N, Bull PM, Landgraf R, Dayanithi G, Leng G. Intracellular calcium stores regulate activity-dependent neuropeptide release from dendrites. Nature. 2002;418:616-625.

68. Ludwig M, Bull PM, Tobin VA, et al. Regulation of activity-dependent dendritic vasopressin release from rat supraoptic neurones. J Physiol. 2005;564:515-522.

69. Loiz M, Sarenac O, Murphy D, Vasopressin J-Z. Central autonomic control and blood pressure regulation. Curr Hypertens Rep. 2018;20:11.

70. Ludwig M, Apps D, Menzies J, Patel JC, Rice ME. Dendritic release of neurotransmitters. Compr Physiol. 2016;7:235-252.

71. Eliava M, Melchior M, Knobloch-Bollmann HS, et al. A new population of parvocellular oxytocin neurons controlling magnocellular neuron activity and inflammatory pain processing. Neuron. 2016;89:1291-2130.
72. Wahis J, Baudon A, Althammer F, et al. Astrocytes mediate the effect of oxytocin in the central amygdala on neuronal activity and affective states in rodents. *Nat Neurosci*. 2021;24(4):529-541. https://doi.org/10.1038/s41593-021-00800-0

73. Hernández-Pérez OR, Hernández VS, Nava-Kopp AT, et al. A synaptically connected hypothalamic magnocellular vasopressin-locus coeruleus neuronal circuit and its plasticity in response to emotional and physiological stress. *Front Neurosci*. 2019;13:196.

74. Zhang B, Qiu L, Xiao W, et al. Reconstruction of the hypothalamo-neurohypophysial system and functional dissection of magnocellular oxytocin neurons in the brain. *Neuron*. 2021;109:331-346.

75. Tang Y, Benusiglio D, Lefevre A, et al. Social touch promotes interfemale communication via activation of parvocellular oxytocin neurons. *Nat Neurosci*. 2020;23:1125-1137.

76. Hasan MT, Althammer F, Silva da Gouveia M, et al. A fear memory engram and its plasticity in the hypothalamic oxytocin system. *Neuron*. 2019;103:133-146.

77. Zhang L, Hernández VS, Vázquez-Juárez E, Chay FK, Barrio RA. Thirst is associated with suppression of habenula output and active stress coping: is there a role for a non-canonical vasopressin-glutamate pathway? *Front Neural Circuits*. 2016;10:13.

78. Jiang-Xie LF, Yin L, Zhao S, et al. A common neuroendocrine substrate for diverse general anesthetics and sleep. *Neuron*. 2019;102:1053-1065.

79. Tomczak A, Mortensen JM, Winneburg R, et al. Interpretation of biological experiments changes with evolution of the Gene Ontology and its annotations. *Sci Rep*. 2018;8:5115.

80. Nguyen TM, Shafi A, Nguyen T, Draghici S. Identifying significantly impacted pathways: a comprehensive review and assessment. *Genome Biol*. 2019;20:203. https://advaitabio.com/pathwayguide/pathway-analysis-vs-gene-set-analysis/

**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Pauža AG, Mecawi AS, Paterson A, et al. Osmoregulation of the transcriptome of the hypothalamic supraoptic nucleus: A resource for the community. *J Neuroendocrinol*. 2021;33:e13007. https://doi.org/10.1111/jne.13007