A Clock-Driven Neural Network Critical for Arousal

Benjamin J. Bell[1,2]*, Qiang Liu[2]*, Dong Won Kim[3], Sangsoo Lee[2], Qili Liu[2], Ian Blum[2], Annette Wang[2], Joe Bedont[3], Anna Chang[3], Habon Issa[2], Jeremiah Cohen[3], Seth Blackshaw[3], and Mark N. Wu[2,3,#]

[1]McKusick-Nathans Department of Genetic Medicine, Johns Hopkins University, Baltimore, MD 21287

[2]Department of Neurology, Johns Hopkins University, Baltimore, MD 21205

[3]Department of Neuroscience, Johns Hopkins University, Baltimore, MD 21205

*These authors contributed equally.

[#]Correspondence should be addressed to M.N.W. (marknwu@jhmi.edu)

Summary

The daily cycling of sleep and arousal states are among the most prominent biological rhythms under circadian control. While much is known about the core circadian clock[1,2], how this clock regulates sleep and arousal remains poorly understood[3]. In *Drosophila*, we previously characterized WIDE AWAKE (WAKE), a clock-output molecule that promotes sleep at night[4,5]. Here, we show that the function of WAKE is conserved in mice and that mWAKE$^+$ neurons comprise a clock-regulated arousal system. Besides the suprachiasmatic nucleus (SCN), mWAKE is expressed in the dorsomedial hypothalamus (DMH) and other arousal-related regions. *mWAKE* mutants exhibit increased spiking of mWAKE$^+$ neurons and increased behavioural arousal at night. In the DMH, mWAKE labels a novel circuit that promotes wakefulness and is required for rapid eye-movement sleep. Single-cell sequencing, immunostaining, and patch-clamp experiments reveal that mWAKE$^+$ neurons constitute a glutamatergic/gabaergic population that projects widely, receives neuromodulatory input, and acts on neuromodulatory neurons. Strikingly, broad chemogenetic silencing of mWAKE$^+$ cells leads to profound loss of behavioural responsiveness and low amplitude, low frequency electroencephalography waveforms. These findings suggest that the genetic mechanisms regulating circadian control of sleep and arousal are conserved across >500 million years of evolution and define a clock-regulated neural network critical for arousal.

Main

The function of WAKE is conserved in mammals

We previously identified the clock-output molecule WIDE AWAKE (WAKE) from a forward genetic screen in *Drosophila*[4]. WAKE modulates the activity of arousal-promoting clock neurons at night, in order to promote sleep onset and quality[4,5]. The mammalian proteome contains a single ortholog, mWAKE (also named ANKFN1/Nmf9), with 56% sequence similarity and which is enriched in the core region of the master circadian pacemaker suprachiasmatic nucleus (SCN)[4,6] (Fig. 1a, Extended Data Fig. 1a). To investigate whether the function of WAKE is conserved in mice, we generated a putative null allele of *mWAKE* (*mWAKE*$^{(-)}$) by CRISPR/Cas9 insertion of 8 base pairs (containing a stop codon and generating a downstream frameshift) in exon 4, which is predicted to be in all splice isoforms of mWAKE (Fig. 1b). As expected, *mWAKE* expression, as assessed by quantitative PCR and *in situ* hybridization (ISH), was markedly reduced in *mWAKE*$^{(-/-)}$ mice, likely due to nonsense-mediated decay (Fig. 1c, 1d). Given *mWAKE* expression in the SCN, we first examined locomotor circadian rhythms and found that *mWAKE*$^{(-/-)}$ mice exhibit a mild but non-significant decrease in circadian period length (Extended Data Fig. 1b, 1c). These results are similar to findings from fly *wake* mutants and mice bearing the *Nmf9* mutation (a previously identified ENU-generated allele of *mWAKE*)[4,6].

Because we previously demonstrated that WAKE mediates circadian regulation of sleep timing and quality in fruit flies[4,5], we next assessed sleep in *mWAKE*$^{(-/-)}$ mice via electroencephalography (EEG). Under light:dark (L:D) conditions, there was no difference in the amount of wakefulness, non-rapid eye movement (NREM), or REM sleep between *mWAKE*$^{(-/-)}$ mutants and wild-type (WT) littermate controls (Extended Data Fig. 1d). In constant darkness (D:D), there is a modest main effect of genotype on wakefulness ($P<0.05$) and NREM sleep ($P<0.05$), and a mild but significant decrease in REM sleep in *mWAKE*$^{(-/-)}$ mutants (Fig. 1e). Although the amount of wakefulness did not appreciably differ in *mWAKE*$^{(-/-)}$ mutants compared to controls, there was a change in the distribution of wakefulness at night; mutants spent more daily time in prolonged wake bouts, and some mutants exhibited dramatically long bouts of wakefulness (Extended Data Fig. 1e, 1f).

Because fly WAKE mainly functions at night and mice (as nocturnal animals) are generally awake at that time, we reasoned that arousal, and not sleep, may be primarily affected in *mWAKE* mutants. Alterations in arousal level can be quantified across different parameters, including sleep/wake behaviour, locomotor activity, and responsiveness to sensory stimuli[7]. Thus, we next examined baseline homecage locomotor activity (Extended Data Fig. 2a). *mWAKE*$^{(-/-)}$ mutants were markedly hyperactive during the subjective night compared to littermate controls, although a mild but significant increase in locomotor activity was also noted during the subjective day (Fig. 1f, 1g, Extended Data Video 1). To rule out the possibility of 2nd site mutations causing this phenotype, we examined transheterozygous *mWAKE*$^{(Nmf9/-)}$ mutants, which also demonstrated robust locomotor hyperactivity during the night, but not during the day (Fig. 1f, 1g). Similar data were obtained for *mWAKE*$^{(-/-)}$, *mWAKE*$^{(+/-)}$, and *mWAKE*$^{(Nmf9/-)}$ mice under L:D conditions (Extended Data Fig. 2b, 2c). Locomotor activity for heterozygous *mWAKE*$^{(+/-)}$ mice was not different from littermate controls, and there was a trend for it being reduced, compared to *mWAKE*$^{(-/-)}$ (P=0.13) and *mWAKE*$^{(Nmf9/-)}$ (P=0.06) (Fig. 1f, 1g). The variability of the pronounced nighttime locomotor activity in *mWAKE*$^{(-/-)}$ and *mWAKE*$^{(Nmf9/-)}$ mutants was driven by intense stereotypic "circling" behaviour in a subset (~30-40%) of these animals (Extended Data Video 1). Related to this, *mWAKE* was previously identified as *Nmf9*, and mutations in this gene were noted to cause circling behaviour, which was interpreted to be due to deficits in vestibular function[6]. However, the intense coordinated circling behaviour demonstrated by *mWAKE* mutants can be seen in hyperaroused mice[8-10]. In addition, in our swim tests, *mWAKE*$^{(-/-)}$ and *mWAKE*$^{(Nmf9/-)}$ mutants swam vigorously without impairment and never had to be rescued from potential drowning, which contrasts with the findings of Zhang et al. (2015) (Extended Data Video 2). Thus, our interpretation is that these and other phenotypes described for *mWAKE*$^{(Nmf9/Nmf9)}$ mutants stem from their hyperarousal, although we cannot rule out subtle vestibular dysfunction requiring specialized testing.

In addition to locomotor activity, another measure of arousal is sensory responsiveness[7]. To characterize this phenotype, we evaluated startle response to an acoustic stimulus in *mWAKE*$^{(-/-)}$ mutants during the subjective day and subjective night. *mWAKE*$^{(-/-)}$ mutants exhibited an increased startle response to 100 dB and 110 dB stimuli during subjective night, but not during the subjective day (Fig. 1h, Extended Data Fig.

2d). To examine provoked locomotor arousal, we performed open-field tests and found that *mWAKE*$^{(-/-)}$ mutants were hyperactive both during the day and night and, unlike controls, failed to demonstrate habituation (Extended Data Fig. 2e-h). Taken together, these data suggest that mWAKE mainly acts to suppress arousal at night, but that specific provoked conditions can also reveal underlying hyperarousal in *mWAKE* mutants during the day.

In *Drosophila*, we previously showed that the large ventrolateral (l-LNv) photorecipient clock neurons in *wake* mutants lose the rhythmicity of their firing at dawn vs dusk. This phenotype results from an increase in spiking frequency specifically at night in the mutants, which is likely due to a reduction in GABA sensitivity in these clock neurons[4]. We thus asked whether loss of mWAKE would cause a similar phenotype in mWAKE^{+} clock neurons from the photorecipient core region of the SCN. To genetically target mWAKE^{+} neurons, while simultaneously creating a mutant allele, we generated transgenic mice where exon 5 was replaced with a tdTomato-p2A-Cre cassette, which should produce a null allele (*mWAKE*$^{(Cre)}$) (Fig. 1i). As predicted, this transgene labels neurons in the core region of the SCN (Fig. 1j), and we further confirmed the overall fidelity of the expression pattern of this transgene, by comparing it to data from whole-brain RNAscope ISH labeling of *mWAKE*[13]. We assessed locomotor activity in D:D in *mWAKE*$^{(Cre/Cre)}$ animals vs heterozygous controls and found that the homozygous animals phenocopied the nighttime hyperactivity of *mWAKE*$^{(-/-)}$ and *mWAKE*$^{(Nmf9/-)}$ mutants, arguing that *mWAKE*$^{(Cre)}$ is a bona fide *mWAKE* mutant allele (Extended Data Fig. 2i, 2j). Patch-clamp recordings reveal a loss of cycling of spiking frequency in mWAKE^{+} SCN neurons, due to a selective increase in spiking frequency at night in *mWAKE*$^{(Cre/Cre)}$ mutants (Fig. 1l, Extended Data Fig. 3a, Table 1). The effect of mWAKE on neuronal excitability is likely cell-autonomous, as intrinsic excitability of mWAKE^{SCN} neurons in these mutants was increased during the night, but not the day (Extended Data Fig. 3b, 3c). Also analogous to fly *wake* mutants, recordings from mWAKE^{SCN} neurons revealed a decrease in GABA-evoked current at night in *mWAKE*$^{(Cre/Cre)}$ mice (Fig. 1m, 1n). This phenotype was also observed in mutants lacking the core clock protein Bmal1[11], suggesting that mWAKE and BMAL act in the same pathway (Extended Data Fig. 3d, 3e). These findings suggest the basic function of WAKE in clock neurons is shared between flies and mice. However, in contrast to fly

WAKE, mWAKE action in the SCN is not crucial for suppressing arousal, because conditional knockout of mWAKE in the ventral forebrain, including the SCN, using the *Six3*^{Cre} driver[12] (*Six3*$^{(Cre/+)}$>*mWAKE*$^{(flox/-)}$ mice) did not alter locomotor activity (Extended Data Fig. 4a-e). Thus, we next sought to identify the brain region(s) where mWAKE acts to regulate arousal.

mWAKE defines a novel clock-regulated arousal circuit in the DMH

Beyond the SCN, mWAKE is expressed in a variety of regions across the brain, such as hypothalamic sub-regions (including the dorsomedial hypothalamus/DMH, Fig. 2a, 2b, also see Fig. 3a), areas implicated in arousal (see below), the limbic system, sensory processing, and limited regions in the cortex[6,13]. The DMH region has previously been implicated in the circadian regulation of arousal[14-16], although the specific neurons involved remain unknown. We thus asked whether mWAKE^{DMH} neurons mediate clock-dependent arousal. Interestingly, in controls, these neurons had greater spiking frequency during the night vs the day, which is antiphase to the SCN, but aligned with the active phase of the nocturnal mouse. Moreover, this nighttime increase was accentuated in *mWAKE*$^{(Cre/Cre)}$ mutants (Fig. 2c, 2d, Extended Data Fig. 5a, Table 1). Similarly, intrinsic excitability of DMH^{mWAKE} neurons was greater during the night, but not the day, in *mWAKE*$^{(Cre/Cre)}$ mutants (Extended Data Fig. 5b, 5c). These data are consistent with a role for mWAKE^{DMH} neurons as circadian-dependent arousal-promoting neurons.

To address the function of mWAKE in the DMH, we performed conditional knockout of mWAKE, by stereotaxically injecting an AAV viral vector expressing Cre-recombinase (AAV-Cre) into the DMH of *mWAKE*$^{(flox/flox)}$ mice (Fig. 2e, 2f, Extended Data Table 2). As shown in Fig. 2g, reduction of mWAKE in the DMH led to a significant increase in nighttime, but not daytime, locomotor activity, compared to their baseline or sham controls. These findings, coupled with the observation that *mWAKE* mutants demonstrate increased spiking of mWAKE^{DMH} neurons at night (Fig. 2c, 2d), suggested that these neurons promote arousal or wakefulness. To test this possibility, we chemogenetically activated these neurons by injecting an AAV vector carrying Cre-dependent DREADD-hM3Dq (AAV-DIO-DREADD-Gq)[17] into the DMH of *mWAKE*$^{(Cre/+)}$ mice (Fig. 2h, 2i, Extended Data Table 2). CNO-mediated activation of mWAKE^{DMH} neurons resulted in a substantial increase in locomotor activity, compared

to vehicle-treated animals (Fig. 2j). Moreover, chemogenetic activation of mWAKE^{DMH} neurons markedly increased wakefulness, with concomitant reductions in NREM and REM sleep (Fig. 2k, 2l, Extended Data Fig. 5d, 5e). We next performed chemogenetic inhibition of mWAKE^{DMH} neurons, by injecting an AAV vector encoding Cre-dependent DREADD-hM4Di (AAV-DIO-DREADD-Gi) and found no significant effects on locomotor activity or amount of wakefulness or NREM sleep (Extended Data Fig. 5f-i, Table 2). However, the amount of REM sleep was significantly reduced following injection of CNO, compared to vehicle alone (Extended Data Fig. 5j). Taken together, these data suggest that mWAKE^{DMH} neurons promote arousal and are necessary for REM sleep and that mWAKE acts to reduce arousal at night by inhibiting the activity of these neurons at that time.

To gain insights into the molecular identity of mWAKE^{DMH} neurons, we conducted single-cell RNA sequencing (scRNA-Seq) of FACS-sorted tdTomato^{+} cells from the hypothalami of *mWAKE*$^{(Cre/+)}$ mice (Fig. 3a, Extended Data Fig. 6a, 6b, Table 3). In addition to neurons, there was a significant population of mWAKE^{+} ependymal cells, which were likely overrepresented due to their ability to survive the dissociation process (Extended Data Fig. 6a). The identity of different neuronal mWAKE^{+} clusters and the spatial location of these clusters were determined by comparison with a hypothalamic scRNA-Seq database and using specific gene markers (Extended Data Fig. 6c)[18]. This analysis revealed 11 mWAKE^{+} clusters in the hypothalamus, with a prominent DMH cluster and 5 clusters localized to the SCN (Fig. 3a). Collectively, mWAKE^{+} neurons comprised a heterogeneous group, but were largely GABAergic (most clusters including all SCN clusters) or glutamatergic (DMH and VMH) in nature (Fig. 3b). We confirmed this observation using RNAscope ISH for the SCN and DMH (Extended Data Fig. 6d-g).

Historically, investigations of the neural basis of arousal have focused on neuromodulatory circuits, but there is growing recognition that GABA/glutamatergic neurons likely form the core substrate for arousal, which is tuned by neuromodulatory networks[3,19,20]. We thus used our scRNA-Seq dataset to characterize the repertoire of neuromodulatory receptors in mWAKE^{+} neurons across hypothalamic sub-groups, and particularly in the DMH (Fig. 3c, Extended Data Fig. 6h). mWAKE^{DMH} neurons express specific noradrenergic, cholinergic, and orexinergic receptors, and so we conducted

rabies virus retrograde tracing (Extended Data Table 2) from these neurons to evaluate potential inputs (Extended Data Fig. 7a). We observed significant retrograde labeling of HDC^+ neurons in the tuberomammillary nucleus (TMN) and orexinergic neurons in the lateral hypothalamus (LH) (Extended Data Fig. 7b, 7c). We also found some labeling of cholinergic neurons in the basal forebrain (BF) and rare labeling of noradrenergic neurons in the locus coeruleus (LC) (Extended Data Fig. 7d, 7e).

To address whether $mWAKE^{DMH}$ neurons functionally respond to these neuromodulators, we performed whole-cell patch-clamp recordings of these neurons from brain slices of *mWAKE*$^{(Cre/+)}$ mice in the presence and absence of norepinephrine, orexin, acetylcholine, or histamine (Fig. 3d, Extended Data Fig. 7f-h, Table 1). Application of orexin and acetylcholine resulted in a significant elevation in spontaneous firing rate of $mWAKE^{DMH}$ neurons in a cell-autonomous manner (i.e., in the presence of synaptic blockers). Norepinephrine increased $mWAKE^{DMH}$ neuron spiking indirectly, consistent with the relative paucity of noradrenergic neurons labeled by retrograde tracing from $mWAKE^{DMH}$ neurons (Extended Data Fig. 7e). In contrast, administration of histamine directly reduced $mWAKE^{DMH}$ neuron firing rate (Extended Data Fig. 7f-h), which may reflect a relative enrichment of the inhibitory H3 receptor subtype in these neurons (Fig. 3c).

Neurons that comprise arousal-promoting nuclei can either act as local interneurons or project broadly to influence the activity of multiple brain regions, including the neocortex[21]. To assess which of these categories $mWAKE^{DMH}$ neurons belong to, we injected an AAV vector expressing Cre-dependent eGFP (AAV-FLEX-eGFP) into the DMH of *mWAKE*$^{(Cre/+)}$ mice and imaged the GFP^+ projections (Fig. 3g, Extended Data Table 2). As shown in Fig. 3h, $mWAKE^{DMH}$ neurons sent projections throughout the brain, including the BF, caudate/putamen, a white matter tract bordering the neocortex, and regions in the brainstem including the LC. Prior work had suggested that the DMH region may mediate circadian timing of arousal by modulating LC firing, although the responsible neurons in the DMH were not identified[15]. Thus, we asked whether $mWAKE^{DMH}$ neurons can regulate the activity of noradrenergic LC neurons. We injected an AAV vector encoding a Cre-dependent Channelrhodopsin2 (AAV-DIO-hChrR2) into the DMH of *mWAKE*$^{(Cre/+)}$*;TH-GFP* mice and then performed whole-cell patch-clamp recordings from noradrenergic TH^{LC} neurons following blue-light

stimulation of the terminals of mWAKE^{DMH} neurons in the LC (Fig. 3i, Extended Data Table 2). In the majority of cases, optogenetic activation of mWAKE^{DMH} triggered noradrenergic TH^{LC} excitatory post-synaptic currents (EPSCs) and spiking (Fig. 3j-l). In summary, our findings suggest that glutamatergic mWAKE^{DMH} neurons are arousal-promoting, project widely, and bidirectionally interact with neuromodulatory networks.

The mWAKE^{+} network intersects with multiple arousal-related regions and is critical for arousal

In addition to the DMH, mWAKE is expressed in or near several classical arousal centers[13]. Using tdTomato labeling from *mWAKE$^{(Cre/+)}$* mice, we observed mWAKE^{+} neuronal cell bodies and projections in the BF, the LH, TMN, in a subregion of the dorsal raphe (DR), and adjacent to the LC (Fig. 3a-d). As expected, these mWAKE^{+} cells do not express markers of the relevant neuromodulator (Extended Data Fig. 8a-e), but instead are largely glutamatergic or GABAergic (Extended Data Fig. 8f-i, Fig. 3a, 3b). Similar to mWAKE^{DMH} neurons, we found that mWAKE^{DR} and mWAKE^{LC} neurons send diffuse projections to the BF, hypothalamus, and brainstem structures (Fig. 4e-h, Extended Data Table 2).

Given these findings, coupled with the expression and function of mWAKE in the DMH, we hypothesized that mWAKE may define a distributed arousal network. Because silencing mWAKE^{DMH} neurons led to a mild phenotype, we chose to inhibit the broad mWAKE^{+} network. To do this, we crossed transgenic mice expressing Cre-dependent DREADD-hM4Di (*LSL-Gi*)[22] to *mWAKE$^{(Cre/+)}$* mice to generate *mWAKE$^{(Cre/+)}$; LSL-Gi* progeny. We first examined the expression pattern of the DREADD-hM4Di in these mice by immunostaining and found that it largely recapitulated the original expression pattern of the *mWAKE$^{(Cre/+)}$* mice (Extended Data Fig. 8j). We next characterized behavioural and EEG phenotypes of these mice. CNO treatment of *mWAKE$^{(Cre/+)}$;LSL-Gi* mice resulted in a profound decrease in arousal level, with reduced or minimal responsiveness to gentle touch or acoustic stimuli, but maintenance of righting reflex (Fig. 4i, 4j; Extended Data Video 3). Strikingly, EEG analyses of these mice revealed a marked shift towards low amplitude, slow frequency waveforms following injection of CNO, but not vehicle control (Fig. 4k-n). For all but one animal, both the behavioural and EEG phenotypes were reproducible and reversible, becoming indistinguishable from

vehicle-injected animals after 24 hrs. However, one CNO-treated animal died after >24 hrs. These phenotypes were suggestive of a stupor-like state, rather than sleep, and indeed *mWAKE*$^{(Cre/+)}$*;LSL-Gi* mice appeared to exhibit a rebound of slow-wave oscillatory activity after the effects of the CNO dissipated (Fig. 4k). For comparison, we assessed the behavioural and EEG effects of chemogenetically silencing arousal-promoting HDC^{+} (histamine decarboxylase) neurons by repeating these experiments with *HDC*$^{(Cre/+)}$*; LSL-Gi* mice. CNO treatment of these mice led to no observable differences in behavioural responsiveness or EEG spectral power and amplitude (Extended Data Fig. 8k-o). These data suggest that the mWAKE^{+} network is essential for basic arousal.

Our studies of WAKE in flies and mice suggest that the core function of WAKE is to reduce arousal at night, by inhibiting the excitability of mWAKE^{+} arousal-promoting neurons at that time[4,5]. Although it is widely accepted that the circadian clock regulates sleep and arousal, the cellular mechanisms remain unclear. In principle, this process could be driven by direct action of SCN projections on arousal circuits[14,15,23], by SCN release of diffusible substances[24-27], or by the activity of local clocks in arousal-promoting nuclei[28]. Our data suggest that mWAKE mediates local clock control of arousal in the DMH, defining the first such neural circuit in this region at the genetic level. Moreover, the observations that WAKE is expressed in or near several arousal-related regions and that these neurons have broad projections suggest a new model for clock control of arousal—multi-focal local control, which may allow for flexible modulatory input from environmental or internal state factors, and yet also facilitate coordinated regulation throughout the brain. Although many nuclei promoting wakefulness have been identified in the mammalian brain[3], very few have been shown to be critical for fundamental arousal. One prominent example is the parabrachial nucleus (PB); large excitotoxic lesions in the PB lead to a coma-like state, with low amplitude, slow oscillations on EEG, a phenotype resembling that seen with silencing the mWAKE^{+} network[29]. For decades, psychologists have proposed the existence of distinct, but reciprocally connected, substrates for arousal: a "higher-order" modulatory system and a basal network, required for attention and consciousness[30]. Our data suggest the possibility that mWAKE marks a basal arousal network that is tuned by neuromodulatory inputs and under circadian clock control.

Methods

Animals

All animal procedures were approved by the Johns Hopkins Institutional Animal Care and Use Committee. All animals were group housed and maintained with standard chow and water available *ad libitum*. Unless otherwise noted, adult male mice (2-4 months) were used in all immunohistochemistry, *in situ* hybridization, and behavioural experiments. All mouse strains were backcrossed at least seven times prior to use in behavioural experiments. Genotyping was performed either by in-house PCR and restriction digest assays, or via Taq-Man based rtPCR probes (Transnetyx). The *mWAKE*^{Nmf9}, *HDC*^{Cre}, *TH-GFP*, and *Six3*^{Cre} mice were obtained from B. Hamilton (University of California, San Diego), A. Jackson (University of Connecticut), A. McCallion (Johns Hopkins University), and Y. Furuta (Memorial Sloan Kettering Cancer Center), respectively. *Bmal1*$^{-}$ (stock number: 009100), and *LSL-Gi* (stock number: 026219) mice were obtained from The Jackson Laboratory.

The *mWAKE* null mutant allele (*mWAKE*$^{-}$) was generated by CRISPR/Cas9 genome editing (Johns Hopkins Murine Mutagenesis Core), using a targeting guide RNA (gRNA: CGC AGA AGA ATC CTC GCA AT) to the 4th exon of *mWAKE* and a 136 bp oligonucleotide (AGT GCG GAC TTT CTC TGG CTC CTG TCC GCA GAA GAA TCC TCG GCG GAA TTC AAT GGG CAC GTT GTT GGT CAT GAT GGC GAT GTC CAG GGG TGT CAG CCC TTC GCT GTT CGG TGT) containing two 64 bp homology arms, surrounding an 8 bp insertion (GCGGAATT), which includes an in-frame stop codon and induces downstream frameshifts. These constructs were injected into the pronucleus of *C57BL/6J* fertilized zygotes and implanted into pseudopregnant females. Pups were assayed for insertion by Sanger sequencing of *mWAKE* gDNA, and knockdown confirmed via qPCR and *in situ* hybridization. The conditional *mWAKE* knockout allele (*mWAKE*^{flox}) was generated via homologous recombination in hybrid (*129/SvEv* x *C57BL/6*) mouse embryonic stem cells (Ingenious Targeting Laboratory). In the stem cells, a construct containing two loxP sites and a Neomycin cassette flanking the 4th exon of *mWAKE* was integrated into the genomic DNA. These cells were then injected into *C57BL/6J* blastocysts, and offspring with high agouti content were crossed to flippase (FLP)-expressing *C57BL/6J* to remove the Neomycin selective marker.

Offspring were then screened via sequencing to confirm proper insertion of both LoxP loci. A transgenic mouse line expressing both Cre recombinase and tdTomato in mWAKE^{+} cells (*mWAKE^{Cre}*) was generated via homologous recombination in hybrid (*129/SvEv* x *C57BL/6*) mouse embryonic stem cells (Ingenious Targeting Laboratory). The knock-in vector targeted exon 5 of the *mWAKE* locus and was integrated 21 bp into exon 5, replacing the remainder of the exon with a *tdTomato-P2A-split Cre-Neo-WPRE-BGHpA* cassette, which causes frameshifted nonsense mutations downstream, resulting in an *mWAKE* loss-of-function allele. Neomycin was excised via FLP mice, and offspring sequenced to confirm the inclusion of the whole sequence into the *mWAKE* locus.

Molecular Biology

To quantify *mWAKE* transcript in the *mWAKE$^{(-/-)}$* mutant mice, qPCR was performed. Hypothalami were dissected ~ZT0, and RNA extracted using Trizol Reagent (Invitrogen). qPCR was performed using a SYBR PCR master mix (Applied Biosystems) and a 7900 Real Time PCR system (Applied Biosystems), with the following primers: mWAKE-F: 5'-CCC TAA CGG TCA GCT TTC AAG A-3' and mWAKE-R: 5'-GAC ATG CTC CAT TCC ACT TTG TAC-3'. GAPDH was used as an internal control. Ct value was compared against regression standard curve of the same primers. 3 biological replicates were performed.

Single Cell Sequencing

Seven week old, male *mWake^{Cre}* mice were processed at ~ZT5 for single-cell RNA-Sequencing (scRNA-Seq). A modified Act-Seq[31] method was used in conjunction with a previously described dissociation protocol[18], with supplementation of Actinomycin D during dissociation (45 μM) and after final resuspension (3 μM), following debris removal (Debris Removal Solution (130-109-398, MACS Miltenyi Biotec) in between. 1 mm hypothalami sections between Bregma 0.02 mm (collecting medial and lateral preoptic area) and Bregma -2.92 mm (beginning of the supramammillary nucleus) were collected between 2 and 3 mice per scRNA-Seq library.

Following dissociation, tdTomato^{+} cells were flow-sorted using an Aria IIu Sorter

(Becton Dickinson). Between 400 - 1000 cells were flow-sorted per brain. Flow-sorted cells were pelleted and re-suspended in 47.6 µl resuspension media. 1 µl of flow-sorted tdTomato$^+$ cells were used to quantify % of tdTomato$^+$ cells with a phase-contrast microscope. Only samples containing ~99% flow-sorted tdTomato$^+$ cells were processed for scRNA-Seq. The remaining 46.6 µl were used for the 10x Genomics Chromium Single Cell system (10x Genomics, CA, United States) using V3.0 chemistry per manufacturer's instructions, generating a total of 3 libraries. Libraries were sequenced on Illumina NextSeq 500 with ~150 million reads per library (~200,000 median reads per cell). Sequenced files were processed through the CellRanger pipeline (v 3.1.0, 10x Genomics) using a custom mm10 genome (with *tdTomato-P2A-Cre-WPRE-bGH* sequence). All 3 libraries were aggregated together for downstream analysis.

Seurat V3[32] was used to perform downstream analysis following the standard pipeline, using cells with more than 200 genes and 1000 UMI counts, removing mWake-tdTomato$^+$ ependymal cells and non-mWake-cells (~1% of the total cluster composed of oligodendrocytes and astrocytes) using known markers genes in the initial clustering[18]. Louvain algorithm was used to generate different clusters, and spatial information (spatial location of different mWake-tdTomato$^+$ clusters across hypothalamic nuclei) and identity of neuronal clusters were uncovered by referring to a hypothalamus scRNA-Seq database[18]. Region-specific transcription factors expressed in mWake$^+$ neurons were used to train mWake-scRNA-Seq. Spatial information of different mWake neuronal populations were further validated by matching to the Allen Brain Atlas ISH data using co-coframer[33], as well as matching to known mWake neuronal distribution across the hypothalamus. The percentage of GABAergic (*Slc32a1*$^+$) and Glutamatergic (*Slc17a6*$^+$) neurons within each cluster were calculated.

Receptors of norepinephrine, dopamine, acetylcholine, histamine and orexin were identified in the scRNA-Seq dataset and the normalized gene expression within each cluster was calculated. scRNA-seq data from this study are accessible through GEO Series accession number GSE146166.

Behavioural Analysis

Animals were raised in a common animal facility under a 14:10 hr Light:Dark (LD)

cycle, and then entrained to a 12:12 hr LD cycle for at least 2 weeks before any locomotor or EEG-based behavioural experiments.

Home cage locomotor activity: Animals were separated into new individual cages with access to food and water *ad libitum* and allowed to acclimate for 4 days before data collection. Data were recorded over 2 days of 12:12 hr LD and 2 days of constant darkness (DD) cycles. Locomotor activity was recorded and analyzed using the Opto M3 monitoring system with IR beams spaced 0.5" apart and Oxymax data-acquisition software (Columbus Instruments). Total activity (the total number of beam breaks along the X and Y axis) was measured by recording the number of beams disrupted by the mouse in 10 s intervals. Locomotor activity profiles were generated from the 2nd day of LD or the 1st day of DD.

Wheel-running activity: Adult male mice (3-5 mos) were placed into individual cages with a vertical running wheel (ActiMetrics) and *ad libitum* access to food and water. After 1 week of acclimatization under 12:12 hr LD conditions, 2 weeks of LD wheel-running activity were recorded, followed by 2 weeks of free-running activity in DD. Wheel-running data were acquired in 1 min bins and analyzed using ClockLab software (ActiMetrics). Period estimates were calculated using data from 12 days of DD.

Open Field Test (OFT): OFT were conducted in 9 x 11 inch polyethylene cages using the Opto M3 beam-break setup described above. Auto-Track software (Columbus Instruments) was used to record the X-Y position, distance, and speed of each mouse at 10 Hz frequency. During the 3 hr test, animals were in constant darkness without bedding, food, or water. Each animal was placed in the arena at the indicated time point, and total activity and distance traveled were summed in 5 min bins as readouts. Behavioural data were analyzed using a custom MATLAB (Mathworks) program. Habituation was calculated by summing the total distance of the first 30 min of the trial and comparing it to the total distance of the final 30 min of each trial.

Acoustic Startle: Acoustic startle response was recorded using an SR-LAB Startle Response System (San Diego Instruments) apparatus, which consists of a sound-isolating

cabinet containing a pressure-sensitive plate. Mice were placed into a plexiglass tube (I.D. 5 cm) and then enclosed inside the chamber on the pressure-recording plate for the duration of the trial. Mice were acclimated to the test environment, including 50 dB of background white noise, for 5 minutes before trials began. Each trial consisted of a 20 ms white noise stimulus (100 dB, 110 dB, or 120 dB) presented from a speaker 20 cm above the mouse's head. The response of the animal in the 100 ms afterwards was recorded as vibration intensity on the pressure platform; Vavg was the total activity averaged over the recorded window, while Vmax was the peak response intensity. All three trial tones were repeated 5 times throughout the experiment, in a pseudorandom order and separated by pseudorandom inter-trial intervals (13-17 s). Trials with significant vibration 100 ms before the tone were excluded from the analysis (<5 instances).

Behavioural assessment of reduced arousal: *mWAKE*$^{(Cre/+)}$*;LSL-Gi* mice were assessed 60 min after IP injection of vehicle or 0.3 mg/kg CNO. Behaviour was scored by a blinded investigator, and each mouse classified as "Normal," "Reduced Reactivity," or "Stuporous," based on the individual's spontaneous behaviour and response to gentle handling. "Normal" mice spontaneously explored the environment and briskly responded to stimulation. "Reduced Reactivity" mice were generally immobile, but would try to evade handling. "Stuporous" mice did not react to handling. All animals exhibited righting reflex.

Electroencephalography (EEG)

Surgery: 8-10 week old male mice were anesthetized to surgical depth with a ketamine/xylazine mixture (100 mg/kg and 10 mg/kg, respectively), and all fur was removed from the top of the head. A skin incision was made along the top of the skull in the rostral-caudal direction, and the scalp was cleaned and connective tissue removed. The 3-channel EEG headmount (Pinnacle Technology) was aligned with the front 3 mm anterior to the bregma and glued to the top of the skull. Four guideholes were hand-drilled, and screws inserted to attach the headmount. EMG wires were then inserted into the left and right neck muscles. After skin closing, the headmount was sealed to the skull

using dental cement. All animals recovered from surgery for > 5 days before being affixed to the EEG recording rigs.

EEG Recording: Sleep behavioural data were obtained using the Pinnacle Technology EEG/EMG tethered recording system. Following recovery, animals were placed into an 8 in diameter round acrylic cage with lid, provided *ad libitum* food and water, and tethered to a 100x preamplifier. All mice were housed in a 12:12 hr L:D cycle and acclimated to the cable tethering for ≥5 days prior to recording. EEG and EMG channels were sampled at 400 Hz, high-pass filtered at 0.5 Hz for EEG and 10 Hz for EMG, digitized, and then acquired using Sirena software (Pinnacle Technology).

Analysis: Mouse sleep was scored visually by one or two trained technicians in Sirenia Sleep software (Pinnacle Technology), using raw EEG/EMG traces in 10 s epochs. Each epoch was scored WAKE, NREM, or REM, and epochs with artifacts were marked for exclusion in further analysis. Animals with severe movement artifacts or poor EEG waveforms were excluded from all behavioural datasets. Sleep or wake bouts were identified as >30 s of continuous sleep or wakefulness, respectively. Spectral analysis was performed using custom MATLAB (Mathworks) programs, and all Fast Fourier Transform spectra used 1024 or 512 point size and the Welch's power spectral density estimate. Spectrograms were composed with short-time Fourier transforms with a window size of 30 s, 60% overlap, and smoothened by a rolling Hann window.

Stereotaxic surgeries

8-12 week old male mice were anesthetized to surgical depth with a ketamine/xylazine mixture (100 mg/kg and 10 mg/kg, respectively), and all fur was removed from the top of the head. The mouse was secured into a Stoelting stereotaxis frame, and the microinjector tip was placed on the Bregma and all coordinates zeroed. Small (~0.5 mm) craniotomies were performed to allow for virus injection (50-300 nL at ~25 nl/min). Coordinates, volumes, viruses used and their sources are listed in Extended Data Table 2. Post injection, animals were allowed to heal and express viral genes for ≥ two weeks (for projection and connectivity studies), and ≥ four weeks (for all behaviour and functional manipulations). If sleep behaviour was measured, EEG headmounts were implanted in a

separate surgery. Locations of all viral injections were confirmed by post-hoc immunostaining, and no animals were excluded from our analyses.

Electrophysiological recordings

Male mice between 5-10 weeks old were deeply anesthetized with isoflurane, and the brains quickly removed and dissected in oxygenated (95% O_2, 5% CO_2) ice-cold slicing solution (2.5 mM KCl, 1.25 mM NaH_2PO4, 2 mM $MgSO_4$, 2.5 mM $CaCl_2$, 248 mM sucrose, 26 mM $NaHCO_3$ and 10 mM glucose). Acute coronal brain slices (250 µm) were prepared using a vibratome (VT-1200s, Leica) and then incubated in oxygenated artificial cerebrospinal fluid (ACSF, 124 mM NaCl, 2.5 mM KCl, 1.25 mM NaH_2PO_4, 2 mM $MgSO_4$, 2.5 mM $CaCl_2$, 26 mM $NaHCO_3$ and 10 mM glucose, 290-300 mOsm) at 28°C for 30 min and then at room temperature for 1 hr. Slices were then transferred to a recording chamber, continuously perfused with oxygenated ACSF at room temperature and visualized using an upright microscope (BX51WI, Olympus). Labeled cells of interest were visualized using infrared differential interference contrast (IR-DIC) and native fluorescence. Glass electrodes (5-8 MΩ) were filled with the following internal solution (130 mM K-gluconate, 5 mM NaCl, 10 mM $C_4H_8N_3O_5PNa_2$, 1 mM $MgCl_2$, 0.2 mM EGTA, 10 mM HEPES, 2 mM MgATP and 0.5 mM Na_2GTP, pH 7.2-7.3, 300 mOsm). Whole-cell patch clamp recordings were obtained using a Multiclamp 700B amplifier (Molecular Devices). Data were sampled at 20 kHz, low-pass filtered at 2 kHz, and digitized using a Digidata 1440A (Molecular Devices).

For baseline spontaneous and evoked firing rate measurements, recordings were performed under current clamp configuration. Baseline recordings were performed for at least 30 s to measure spontaneous firing rate. To measure evoked firing rate, current injections from -10 to 100 pA were performed for 600 ms. 0.5% biocytin (wt/wt) was added to the internal solution to label the recorded cell. Slices were fixed post-experiment in 4% PFA overnight, and then incubated with Alexa488-conjugated streptavidin (Invitrogen, 1:2000) for 24 hrs, and then imaged on a Zeiss 800 confocal microscope. For measurement of GABA currents, voltage-clamp recordings at -70 mV were performed at ZT12-14. K-gluconate in the patch-pipette was replaced with $CeCl_2$ for these recordings, and CeOH was used to adjust the pH of the internal solution. 1 mM

GABA was delivered for 5 s using a Picospritzer III (Parker), and GABA-evoked current was recorded. For recordings of mWAKE^{DMH} neurons, ~30% cells exhibited spontaneous firing, and so analyses of evoked responses were restricted to this subset of neurons.

For application of norepinephrine, orexin, acetylcholine, and histamine current-clamp recordings were performed at ZT12-14. Baseline spontaneous activity was recorded for 30 s, and then for another 30 s following continuous application of the neurotransmitter. Compounds were pre-loaded into pulled glass pipettes (3-4 μm I.D. at the tip) and delivered at the recorded cell using a Picospritzer III (Parker). The compounds and concentrations used were as follows: norepinephrine (100 μM), acetylcholine chloride (1 mM), orexin-A (300 nm), and histamine (20 μM). For synaptically isolating recorded cells, ion channel blockers (20 μM CNQX; 50 μM AP5; 10 μM Picrotoxin) were added to the ACSF perfusion. For quantification of firing rates for cells reaching plateau potential, the maximum firing rate observed for any cell following application of the relevant neurotransmitter was used.

For optogenetic experiments, AAV-DIO-hChR2 (Extended Data Table 2) was injected into the DMH of *mWAKE*$^{(Cre/+)}$*;TH-GFP* mice. Acute slices were prepared 3 weeks post-viral injection. Current-clamp and voltage-clamp recordings were performed from norepinephrine-expressing cells in the locus coeruleus (TH^{+}), in the presence of optogenetic activation of ChR2-expressing terminals in the LC. Slices were exposed to blue light (480 nm) from the upper lens for 2 s at different stimulation frequencies (5, 10, 20, and 50 Hz) triggered by the Digidata 1440A.

Immunohistochemistry

Mice were deeply anesthetized with a ketamine/xylazine mixture then fixed by transcardial perfusion with 4% paraformaldehyde (PFA). Brains were subsequently drop-fixed in 4% PFA for 24-48 hr and transferred into 1x PBS before being sectioned at 40 μm thickness using a vibratome (VT1200S, Leica). Free-floating sections were washed in 1x PBS, blocked for 1 hr in blocking buffer (PBS containing 0.25% Triton-X-100 and 5% normal goat serum or normal donkey serum), then incubated with rabbit anti-HDC (Progen, 16045, 1:800), goat anti-ChAT (Millipore, AB144P, 1:250), rabbit anti-

TPH2 (Novus, NB100-74555, 1:300), rabbit anti-orexin A (Abcam, AB6214, 1:500), chicken anti-TH (Abcam, AB76442, 1:1000), rat anti-HA (Roche 11867423001, 1:250) or mouse anti-Cre (Millipore, MAB3120, 1:500) in blocking buffer at 4°C overnight. The following day, slices were washed with PBST (PBS and 0.1% Tween-20), then incubated with Alexa 488 anti-rabbit (ThermoFisher, A-11008, 1:2500-1:5000), Alexa 568 anti rabbit (ThermoFisher, A-11011, 1:2500), Alexa 568 anti-goat (ThermoFisher, A11057, 1:2000), Alexa 488 anti-rat (ThermoFisher A-11006, 1:2000), or Alexa 488 anti-chicken (ThermoFisher, A-11039, 1:2000) secondary antibodies for 2.5 hrs in blocking buffer. Brain sections were then washed in PBST, incubated in DAPI (1:2000, Millipore) for 5 min, then washed in PBS. Sections were mounted on slides using VectaShield HardSet Mounting Medium (Vector Laboratories, USA). In all tdTomato images, we visualized native fluorescence of tdTomato. Images were acquired using a Zeiss LSM800 confocal microscope under 10x-63x magnification.

In situ **hybridization**

ISH for *mWAKE* was performed as previously described[4]. RNAScope ISH was performed using the RNAscope 2.5 Chromogenic Assay and the BaseScope™ Detection Reagent Kit according to the manufacturer's instructions (Advanced Cell Diagnostics (ACD))[34]. Target probes were designed to exon 4 of *mWake*. Mouse brains were dissected and fresh frozen in Tissue-Tek O.C.T. (VWR) and cryosectioned at 10 µm. Sections were treated with hydrogen peroxide and protease, before hybridization to the custom *mWAKE* probe and subsequent amplification. Signal was detected by chromogenic reaction with BaseScope™ Fast RED, and sections counterstained with hematoxylin. Images were acquired on a Keyence BZ-X700 microscope (Keyence) under 10x brightfield illumination. Fluorescent *in situ* hybridization (FISH) was performed with the RNAscope™ Fluorescent Multiplex Assay (ACD), using probes targeting *tdTomato* and *Slc32a1 (VGat)* or *Slc17a6 (VGlut2)* mRNA. Preparation of tissue sections was performed as above, followed by simultaneous hybridization to both probes. Probe binding was indicated by deposition of target-specific fluorophores at each location via TSA Plus Fluorescence kit (Perkin-Elmer), and sections were then

counterstained with DAPI. Sections were imaged on a Zeiss LSM880 confocal at 20x and 63x.

Designer Receptors Exclusively Activated by Designer Drugs (DREADDs)

DREADD receptors coupled to either Gq or Gi were expressed in a Cre-dependent fashion in *mWAKE*$^{+}$ neurons of *mWAKE*$^{(Cre/+)}$ mice, either locally in the DMH, via stereotaxic injection of a viral vector (AAV-DIO-DREADD-Gq or AAV-DIO-DREADD-Gi) (Extended Data Table 2), or globally via crossing with a transgenic effector mouse line (*B6N;129-CAG-LSL-HA-hM4Di-mCitrine*, "LSL-Gi"). Clozapine *N*-oxide (CNO) (SigmaAldrich) was prepared as a stock solution of 50 mg/ml in DMSO, and then freshly diluted to 0.1 mg/ml in sterile PBS before IP injection. Solution clarity was monitored throughout dosing, and the solution was warmed to 37°C if precipitates were observed. Vehicle control was prepared as sterile saline + 0.01% DMSO. All injections occurred at the same ZT/CT time within each experiment, and all animals were treated with vehicle or CNO each day in a cross-over design, with ≥2 days recovery between experimental recording days. Each DREADD experiment included CNO controls (high dose 3 mg/kg CNO IP injected into WT littermates), as well as genetic controls (CNO injections in *mWAKE*$^{(Cre/+)}$ or *LSL-Gi* alone animals).

Statistical analysis

Statistical analyses were performed in Prism 7 and 8 (Graphpad). For comparisons of two groups of normally distributed data, unpaired Student t-tests were used; if these comparisons were before and after treatment of the same animals or cells, paired t-tests were used instead. For comparisons of two groups of non-normally distributed data, Mann Whitney U tests were performed, with a Holm-Bonferroni correction, if required for multiple comparisons. For multiple comparisons of normally distributed data with 2 factors, 2 way ANOVAs were performed (with repeated measures, if applicable), followed by post-hoc Sidak tests. For multiple comparisons of non-normally distributed data, Kruskall-Wallis tests were performed with post-hoc Dunn's tests.

References

1 Allada, R., Emery, P., Takahashi, J. S. & Rosbash, M. Stopping time: the genetics of fly and mouse circadian clocks. *Annual review of neuroscience* **24**, 1091-1119 (2001).

2 Mohawk, J. A., Green, C. B. & Takahashi, J. S. Central and peripheral circadian clocks in mammals. *Annual review of neuroscience* **35**, 445-462 (2012).

3 Scammell, T. E., Arrigoni, E. & Lipton, J. O. Neural Circuitry of Wakefulness and Sleep. *Neuron* **93**, 747-765, doi:10.1016/j.neuron.2017.01.014 (2017).

4 Liu, S. *et al.* WIDE AWAKE mediates the circadian timing of sleep onset. *Neuron* **82**, 151-166, doi:10.1016/j.neuron.2014.01.040 (2014).

5 Tabuchi, M. *et al.* Clock-Generated Temporal Codes Determine Synaptic Plasticity to Control Sleep. *Cell* **175**, 1213-1227 e1218, doi:10.1016/j.cell.2018.09.016 (2018).

6 Zhang, S. *et al.* Nmf9 Encodes a Highly Conserved Protein Important to Neurological Function in Mice and Flies. *PLoS Genet* **11**, e1005344, doi:10.1371/journal.pgen.1005344 (2015).

7 Pfaff, D., Ribeiro, A., Matthews, J. & Kow, L. M. Concepts and mechanisms of generalized central nervous system arousal. *Ann N Y Acad Sci* **1129**, 11-25, doi:10.1196/annals.1417.019 (2008).

8 Homanics, G. E. *et al.* Mice devoid of gamma-aminobutyrate type A receptor beta3 subunit have epilepsy, cleft palate, and hypersensitive behavior. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 4143-4148, doi:10.1073/pnas.94.8.4143 (1997).

9 Glick, S. D., Zimmerberg, B. & Greenstein, S. Individual differences among mice in normal and amphetamine-enhanced locomotor activity: relationship to behavioral indices of striatal asymmetry. *Brain research* **105**, 362-364, doi:10.1016/0006-8993(76)90436-4 (1976).

10 Tilley, M. R. & Gu, H. H. Dopamine transporter inhibition is required for cocaine-induced stereotypy. *Neuroreport* **19**, 1137-1140, doi:10.1097/WNR.0b013e3283063183 (2008).

11 Bunger, M. K. *et al.* Mop3 is an essential component of the master circadian pacemaker in mammals. *Cell* **103**, 1009-1017, doi:10.1016/s0092-8674(00)00205-1 (2000).

12 Furuta, Y., Lagutin, O., Hogan, B. L. & Oliver, G. C. Retina- and ventral forebrain-specific Cre recombinase activity in transgenic mice. *Genesis* **26**, 130-132 (2000).

13 Wang, A., Bell, B., Kim, T., Blackshaw, S. & Wu, M. N. mWAKE is Expressed in Arousal-Related Nuclei and Reveals Novel Brain Regions. *(In preparation)* (2020).

14 Chou, T. C. *et al.* Critical role of dorsomedial hypothalamic nucleus in a wide range of behavioral circadian rhythms. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **23**, 10691-10702 (2003).

15 Aston-Jones, G., Chen, S., Zhu, Y. & Oshinsky, M. L. A neural circuit for circadian regulation of arousal. *Nature neuroscience* **4**, 732-738, doi:10.1038/89522 (2001).

16 Deurveilher, S. & Semba, K. Indirect projections from the suprachiasmatic nucleus to major arousal-promoting cell groups in rat: implications for the

circadian control of behavioural state. *Neuroscience* **130**, 165-183, doi:10.1016/j.neuroscience.2004.08.030 (2005).
17 Krashes, M. J. *et al.* Rapid, reversible activation of AgRP neurons drives feeding behavior in mice. *J Clin Invest* **121**, 1424-1428, doi:10.1172/JCI46229 (2011).
18 Kim, D. W. *et al.* Single cell RNA-Seq analysis identifies molecular mechanisms controlling hypothalamic patterning and differentiation. *bioRxiv* **doi: https://doi.org/10.1101/657148** (2020).
19 Saper, C. B. & Fuller, P. M. Wake-sleep circuitry: an overview. *Curr Opin Neurobiol* **44**, 186-192, doi:10.1016/j.conb.2017.03.021 (2017).
20 Eban-Rothschild, A., Appelbaum, L. & de Lecea, L. Neuronal Mechanisms for Sleep/Wake Regulation and Modulatory Drive. *Neuropsychopharmacology* **43**, 937-952, doi:10.1038/npp.2017.294 (2018).
21 Jones, B. E. Arousal and sleep circuits. *Neuropsychopharmacology* **45**, 6-20, doi:10.1038/s41386-019-0444-2 (2020).
22 Zhu, H. *et al.* Cre-dependent DREADD (Designer Receptors Exclusively Activated by Designer Drugs) mice. *Genesis* **54**, 439-446, doi:10.1002/dvg.22949 (2016).
23 Mistlberger, R. E. Circadian regulation of sleep in mammals: role of the suprachiasmatic nucleus. *Brain Res Brain Res Rev* **49**, 429-454 (2005).
24 Cheng, M. Y. *et al.* Prokineticin 2 transmits the behavioural circadian rhythm of the suprachiasmatic nucleus. *Nature* **417**, 405-410, doi:Doi 10.1038/417405a (2002).
25 Kramer, A. *et al.* Regulation of daily locomotor activity and sleep by hypothalamic EGF receptor signaling. *Science* **294**, 2511-2515 (2001).
26 Kraves, S. & Weitz, C. J. A role for cardiotrophin-like cytokine in the circadian control of mammalian locomotor activity. *Nature neuroscience* **9**, 212-219, doi:10.1038/nn1633 (2006).
27 Pevet, P. & Challet, E. Melatonin: both master clock output and internal time-giver in the circadian clocks network. *J Physiol Paris* **105**, 170-182, doi:10.1016/j.jphysparis.2011.07.001 (2011).
28 Yu, X. *et al.* Circadian factor BMAL1 in histaminergic neurons regulates sleep architecture. *Current biology : CB* **24**, 2838-2844, doi:10.1016/j.cub.2014.10.019 (2014).
29 Fuller, P. M., Sherman, D., Pedersen, N. P., Saper, C. B. & Lu, J. Reassessment of the structural basis of the ascending arousal system. *The Journal of comparative neurology* **519**, 933-956, doi:10.1002/cne.22559 (2011).
30 Coull, J. T. Neural correlates of attention and arousal: insights from electrophysiology, functional neuroimaging and psychopharmacology. *Prog Neurobiol* **55**, 343-361, doi:10.1016/s0301-0082(98)00011-2 (1998).
31 Wu, Y. E., Pan, L., Zuo, Y., Li, X. & Hong, W. Detecting Activated Cell Populations Using Single-Cell RNA-Seq. *Neuron* **96**, 313-329 e316, doi:10.1016/j.neuron.2017.09.026 (2017).
32 Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902 e1821, doi:10.1016/j.cell.2019.05.031 (2019).
33 Lein, E. S. *et al.* Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168-176, doi:10.1038/nature05453 (2007).

34 Wang, F. *et al.* RNAscope: a novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. *J Mol Diagn* **14**, 22-29, doi:10.1016/j.jmoldx.2011.08.002 (2012).

Acknowledgements

We thank A. Meredith and B. McNally for advice on SCN patch-clamp recordings and S. Hattar for assistance with EEG recordings. We thank B. Hamilton, A. Jackson, A. McCallion, and Y. Furuta for sharing mouse strains. We thank the Transcriptomics and Deep Sequencing Core for scRNA sequencing and the Ross Flow Cytometry Core flow sorting. We thank members of the Wu Lab for discussion. This work was supported by NIH grants R01NS094571 and R01NS079584. (M.N.W.) and a NINDS Center grant NS05027 for machine shop work.

Author Contributions

B.B. and M.N.W. conceived the project, with input from J.C. and S.B. B.B. performed most or all mouse genetics, behavioural experiments, EEG recordings, expression analyses, viral injections, and rabies tracing. Q.L.* performed all electrophysiology and optogenetic experiments and assisted with EEG analyses. D.W.K. performed scRNA-Seq analysis. S.L. performed viral injections and immunostaining. Q.L. assisted with mouse genetics and generation of the *mWAKE*$^{-}$ allele. I.B. assisted with EEG analyses. A.W. and J.B. assisted with expression analyses. A.C. assisted with viral injections. H.I. assisted with mouse genetics. B.B. and M.N.W. wrote the manuscript, with feedback from all authors.

Figure Legends

Figure 1 | mWAKE inhibits arousal and SCN firing at night

a, *mWAKE* mRNA detected by BaseScope ISH demonstrates expression in the SCN core (solid and dashed lines denote SCN and core region, respectively). Higher magnification inset shows representative *mWAKE* mRNA expression in cells. Scale bar applies to both images and denotes 200 µm and 50 µm for the entire image and inset, respectively. **b,** Schematic showing genomic structure of the *mWAKE* locus and CRISPR/Cas9-mediated insertion of an in-frame stop codon in exon 4 in the $mWAKE^{(-)}$ mutation. **c,** Relative mRNA level for *mWAKE*, determined by qPCR, in $mWAKE^{(-/-)}$ vs WT littermate control hypothalami (normalized to 1.0) (n=3 trials). **d,** Representative image of ISH using a digoxigenin-labeled *mWAKE* RNA probe in the SCN of $mWAKE^{(-/-)}$ vs WT mice (n=4 for both). Scale bar denotes 500 µm. **e**, Vigilance state (% per 2 hr bin) determined by EEG recordings for $mWAKE^{(-/-)}$ (n=8) vs WT littermate control (n=6) mice under DD conditions. **f**, Profile of locomotor activity (defined by beam-breaks) over 24 hrs for $mWAKE^{(+/+)}$ (n=19, grey), $mWAKE^{(+/-)}$ (n=22. black), $mWAKE^{(-/-)}$ (n=19, red), and $mWAKE^{(Nmf9/-)}$ (n=9, cyan) mice under DD conditions. **g**, Total locomotor activity from CT0-12 and CT12-24 from the mice described in (**f**). **h**, Startle response (Vavg) measured in the first 100 ms following a 100, 110, or 120 dB tone for $mWAKE^{(+/+)}$ (grey, n=10) vs $mWAKE^{(-/-)}$ mice (red, n=10) at CT12-14. The average of 5 responses is shown. **i**, Schematic showing genomic structure of the *mWAKE* locus and replacement of exon 5 with a *tdTomato-p2A-Cre* cassette in the $mWAKE^{(Cre)}$ line. **j**, Native tdTomato fluorescence in the SCN of a $mWAKE^{(Cre/+)}$ mouse (solid and dashed lines denote SCN and core region, respectively). Scale bar represents 200 µm. **k**, Representative membrane potential traces from whole-cell patch clamp recordings of $mWAKE^{SCN}$ neurons from $mWAKE^{(Cre/+)}$ (grey, top) and $mWAKE^{(Cre/Cre)}$ (red, bottom) slices at ZT0-2 and ZT12-14. **l**, Spontaneous mean firing rate for $mWAKE^{+}$ SCN neurons at ZT0-2 and ZT12-14 from $mWAKE^{(Cre/+)}$ (n=21 and n=22) vs $mWAKE^{(Cre/Cre)}$ (n=24 and n=20) mice. **m**, Representative traces of voltage-clamp recordings of $mWAKE^{SCN}$ neurons of $mWAKE^{(Cre/+)}$ vs $mWAKE^{(Cre/Cre)}$ at ZT12-14. Timing of GABA (1 mM) application is shown. **n**, GABA-evoked current in $mWAKE^{SCN}$ neurons of $mWAKE^{(Cre/+)}$ (n=18) vs $mWAKE^{(Cre/Cre)}$ (n=20) at ZT12-14. For panels (**l**) and (**n**), n represents individual cells

from 4 animals for each condition. In this figure and throughout, error bars represent SEM and "*", "**", and "***" denote $P<0.05$, $P<0.01$, and $P<0.001$, respectively.

Figure 2 | mWAKE^{DMH} neuronal activation is sufficient for arousal and required for REM sleep.

a, *mWAKE* mRNA detected by Basescope ISH reveals expression in the dorsomedial hypothalamus (DMH). Dashed lines indicate DMH region, and inset indicates shows representative *mWAKE* mRNA expression in cells. Scale bar applies to both images and denotes 200 µm and 50 µm for the entire image and inset, respectively. **b**, Native tdTomato fluorescence in the DMH of a *mWAKE*$^{(Cre/+)}$ mouse (dashed lines denote DMH region). Scale bar represents 200 µm. **c**, Representative membrane potential traces from whole-cell patch clamp recordings of mWAKE^{DMH} neurons in *mWAKE*$^{(Cre/+)}$ (grey, top) and *mWAKE*$^{(Cre/Cre)}$ (red, bottom) slices at ZT0-2 and ZT12-14. **d**, Spontaneous mean firing rate for mWAKE^{DMH} neurons at ZT0-2 and ZT12-14 from *mWAKE*$^{(Cre/+)}$ (n=19 and n=14) vs *mWAKE*$^{(Cre/Cre)}$ (n=16 and n=12) mice. n represents individual cells from ≥4 animals. **e**, Schematic showing bilateral injections of AAV viral vector containing Cre-recombinase and eGFP (AAV-Cre), or eGFP alone (AAV-Sham) into the DMH of *mWAKE*$^{(flox/flox)}$ mice. **f**, Representative image of eGFP fluorescence expression in the DMH following AAV-Cre injection described in (**e**). Scale bar denotes 200 µm. **g**, Total locomotor activity (total number of beam breaks along X and Y axis) during CT0-4 vs CT12-16 under DD conditions for *mWAKE*$^{(flox/flox)}$ mice before ("pre", grey) and after ("post", green) DMH injection of AAV-Sham (n=6) and AAV-Cre (n=9). **h**, Schematic showing bilateral injections of AAV-DIO-DREADD-Gq into the DMH of *mWAKE*$^{(Cre/+)}$ mice. **i**, Representative image of mCherry fluorescence expression in the DMH of *mWAKE*$^{(Cre/+)}$ following AAV-DIO-DREADD-Gq injection described in (**h**). Scale bar denotes 200 µm. **j**, Total locomotor activity (total number of break breaks along X and Y axis) of *mWAKE*$^{(Cre/+)}$ mice with bilateral injections of AAV-DIO-DREADD-Gq into the DMH in the 4 hrs following IP injection of vehicle vs CNO (1 mg/kg) at CT 8 (n=6). **k**, Representative short-time Fourier transform spectrograms of 8 hrs of recorded EEG activity, starting after IP injection at ZT6 of vehicle alone (above) or 1 mg/kg CNO (below), from *mWAKE*$^{(Cre/+)}$ mice injected with AAV-DIO-DREADD-Gq bilaterally into

the DMH. Power density is represented by the color-scheme and deconvoluted by frequency on the y-axis, over time on the x-axis. **l**, Amount of wakefulness derived from EEG plotted as % time in 1 hr bins following IP injection of vehicle (grey) or 1 mg/kg CNO (red) (n=4) at ZT6.

Figure 3 | Identity of mWAKE$^+$ hypothalamic neurons and their interaction with neuromodulatory systems

a, Schematic (left) showing dissection of hypothalami from *mWAKE*$^{(Cre/+)}$ mice, followed by FACS and scRNA sequencing. UMAP (Uniform Manifold Approximation and Projection) plot (right) showing distribution of mWake$^+$ neurons across hypothalamic nuclei, as determined by single-cell expression profiling. 11 clusters are defined, for SCN, DMH, POA (pre-optic area), TMN (tuberomammillary nucleus), ZI (zona incerta), and VMH (ventromedial hypothalamus) regions. "Gal$^+$" and "Cck$^+$" refer to Galanin$^+$ and Cholecystokinin$^+$. **b**, Bar graph showing distribution of GABAergic and glutamatergic mWAKE$^+$ neurons for each scRNA-Seq neuronal cluster. **c**, Bar graph showing distribution of expression for norepinephrine, dopamine, acetylcholine, histamine and orexin receptors identified in DMH mWAKE$^+$ neurons (cluster 4). **d**, Schematic showing patch-clamp recordings of mWAKE^{DMH} neurons from *mWAKE*$^{(Cre/+)}$ mice, following application of norepinephrine, orexin, or acetylcholine. **e**, Representative membrane potential traces from recordings described in (**d**) following administration of norepinephrine (top, red), orexin (middle, green), or acetylcholine (bottom, blue) in the presence of synaptic blockers at ZT12-14. **f**, Spontaneous mean firing rate of mWAKE^{DMH} neurons following application of 100 μM norepinephrine (n=23 and 12 cells), 300 nM orexin-A (n=8 and 7 cells), or 1 mM acetylcholine chloride (n=8 and 10 cells) at ZT12-14 in the absence ("- synaptic blockers") or presence ("+ synaptic blockers") of 20 μM CNQX, 50 μM AP5, and 10 μM picrotoxin. The neurons recorded from in **f** were derived from 3-6 animals. **g**, Schematic showing unilateral stereotaxic injection of AAV-viral vector expressing Cre-dependent eGFP (AAV-FLEX-eGFP) into the DMH of *mWAKE*$^{(Cre/+)}$ mice. **h**, Representative sagittal section showing broad projection of GFP fluorescence following injection of AAV-FLEX-eGFP shown in (**e**). Dashed ellipse denotes injection site, and locations of the basal forebrain (BF),

cortex (Cx), hypothalamus (Hyp), locus coeruleus (LC), and pons are noted. Scale bar denotes 1 mm. **i**, Schematic showing AAV-DIO-hChR2 injected into the DMH of *mWAKE*$^{(Cre/+)}$*;TH-GFP* mice unilaterally, and whole-cell patch clamp recordings of TH^{LC} neurons, following 480 nm blue-light stimulation. **j** and **k**, Representative excitatory post-synaptic currents (EPSCs) (**j**) or action potentials (APs) (**k**) from TH^{LC} neurons, triggered by optogenetic stimulation of mWAKE^{DMH} neuron terminals at 5 Hz. Blue lines indicate timing of blue-light pulses. **l**, Fidelity (%) of TH^{LC} neuron EPSCs (dark blue) and APs (light blue) triggered by optogenetic stimulation of mWAKE^{DMH} neuron terminals (n=4 cells from 3 mice, described in **i**). A total of 7 TH^{LC} cells from 3 independent virally-injected *mWAKE*$^{(Cre/+)}$*;TH-GFP* mice were recorded, of which 4 exhibited optogenetically-triggered responses.

Figure 4 | mWAKE is expressed in multiple sleep/wake-related regions and labels a network critical for arousal

a-**d**, Representative images of tdTomato fluorescence and anti-Cre labeling of processes and cell bodies of mWAKE^{+} cells in multiple arousal-related regions from *mWAKE*$^{(Cre/+)}$ mice. mWAKE^{+} cell bodies are found in the basal forebrain (BF) (**a**), lateral hypothalamus (LH) (**b**), and tuberomammillary nucleus (TMN), and adjacent to the dorsal raphe (DR) (**c**) and LC (**d**). Insets show anti-Cre labeling in higher magnification images from the regions indicated by the dashed ellipse. LSn, f, Aq, and 4v denote lateral septal nuclei, fornix, cerebral aqueduct, and 4th ventricle, respectively. For inset and main panel, scale bars represent 50 and 400 μm (**a**), 50 and 150 μm (**b**), 50 and 400 μm (**c**), and 50 and 400 μm (**d**). **e** and **g**, Schematic showing unilateral stereotaxic injection of AAV-viral vector expressing Cre-dependent eGFP (AAV-FLEX-eGFP) into the DR (**e**) or LC (**g**) of *mWAKE*$^{(Cre/+)}$ mice. **f** and **h**, Representative sagittal section showing broad projection of GFP fluorescence following injection of AAV-FLEX-eGFP shown in (**e** and **g**). Dashed ellipse denotes injection site, and locations of the BF, hypothalamus (Hyp), DR, LC, and pons are noted. Scale bars denote 1 mm. **i,** Jitter plot showing behavioural classification of arousal level for *mWAKE*$^{(Cre/+)}$*; LSL-Gi* mice following injection of vehicle or 0.3 mg/kg CNO (n=12). **j**, Startle response (Vmax) measured in the first 100 ms following a 120 dB tone for *mWAKE*$^{(Cre/+)}$*; LSL-Gi* mice

following injection of vehicle (grey) or 0.3 mg/kg CNO (magenta) (n=5) at CT12-14. **k**, Representative short-time Fourier transform spectrograms of 12 hrs of recorded EEG activity from *mWAKE*$^{(Cre/+)}$; *LSL-Gi* mice after IP injection at ZT10 of vehicle alone (above) or 0.3 mg/kg CNO (below). **l**, Representative EEG traces for a 30 s window (above) and short-time Fourier transform spectrograms (below) starting 150 mins after injection of vehicle (left) or 0.3 mg/kg CNO (right) at ZT10. The red box indicates the time window used for spectral and amplitude analysis in (**m**) and (**n**). **m,** Welch's power spectral density estimates as a percentage of total EEG power across 10 min, averaged across multiple EEG traces. *mWAKE*$^{(Cre/+)}$*; LSL-Gi* injected with vehicle (grey) or 0.3 mg/kg CNO (blue), n=4. Inset shows a plot of delta-band power as a percentage of total EEG power. **n**, EEG trace amplitude (plotted as normalized root mean square (RMS)) for the animals described in (**m**).

Extended Data Figure 1 | Circadian and EEG-related phenotypes of *mWAKE* mutant mice

a, Schematic showing domain structure of mWAKE in mice and humans, compared to *Drosophila* WAKE. Percentage similarity of mouse and human mWAKE, compared to fly WAKE is shown. **b**, Representative double-plotted actograms of wheel running activity for *mWAKE*$^{(+/+)}$ (left) vs *mWAKE*$^{(-/-)}$ mice (right), covering 14 days of LD then 14 days of free-running in DD. Activity plotted as number of wheel revolutions in 10 min bins. **c**, Locomotor period length (τ) during the DD period for *mWAKE*$^{(+/+)}$ (n=8) and *mWAKE*$^{(-/-)}$ (n=8) mice. **d**, Vigilance state (% per 2 hr bin) determined by EEG recordings for *mWAKE*$^{(-/-)}$ (n=8) vs WT littermate control (n=6) mice under LD conditions. The mice in panel (**d**) are the same as in Fig. 1e. **e**, Hypnograms (*top*) and short-time Fourier transform spectrograms (*bottom*) over 24 hrs in DD showing example of prolonged wake bout in *mWAKE*$^{(-/-)}$ mice, compared to *mWAKE*$^{(+/+)}$ control. Power density is represented by the color-scheme and deconvoluted by frequency on the y-axis, over time on the x-axis. **f**, Wakefulness bout length distribution over a 24 hr period expressed as a % of total wakefulness, for *mWAKE*$^{(+/+)}$ vs *mWAKE*$^{(-/-)}$ mice.

Extended Data Figure 2 | *mWAKE* mutants are hyperactive at night

a, Schematic of the beam-break arena used to measure locomotion, with infrared red beams forming a grid spaced at 0.5" intervals across the floor of the cage. **b**, Profile of locomotor activity (defined by beam breaks) over 24 hrs for *mWAKE*$^{(+/+)}$ (grey, n=19), *mWAKE*$^{(-/-)}$ (red, n=19), and *mWAKE*$^{(Nmf9/-)}$ (cyan, n=10) mice under LD conditions. **c**, Total locomotor activity from the mice in (**b**) at CT0-12 and CT12-24. Note that these data are from the same mice as in Fig. 1f. **d**, Startle response (Vavg) following a 100, 110, or 120 dB tone for *mWAKE*$^{(+/+)}$ (grey, n=10) vs *mWAKE*$^{(-/-)}$ mice (red, n=10) at CT0-2. Note that these data are from the same mice as in Fig. 1h. **e** and **g**, Total distance (cm) traveled in five min bins in an open-field test, across the first 3 hours of the trial (CT0-CT3, **e**) or (CT12-CT15, **g**) for *mWAKE*$^{(+/+)}$ (n=8, grey) vs *mWAKE*$^{(-/-)}$ (n=8, red) mice. Data were collected from 3 independent trials. **f** and **h**, Habituation of animals shown in (**e**) and (**g**), respectively, as assessed by total distance (cm) traveled in the first 30 mins vs the last 30 mins of the trial. **i**, Locomotor activity profile over 24 hrs for *mWAKE*$^{(Cre/+)}$ (grey, n=9) and *mWAKE*$^{(Cre/Cre)}$ (red, n=10) under DD conditions. **j**, Total locomotor activity from the mice in (**e**) at CT0-12 vs CT12-24.

Extended Data Figure 3 | mWAKE inhibits the activity of SCN neurons at night
a, Schematic of whole-cell patch-clamp recordings from mWAKE^{+} neurons (tdTomato^{+} neurons in *mWAKE*$^{(Cre/+)}$ or *mWAKE*$^{(Cre/Cre)}$ mice), in the core region of the SCN (top). Representative images of a biocytin- and tdTomato-labeled cell post-recording (bottom). Scale bar denotes 25 µm. **b** and **c**, *f-I* curves for mWAKE^{+} SCN neurons from *mWAKE*$^{(Cre/+)}$ vs *mWAKE*$^{(Cre/Cre)}$ mice at ZT0-2 (n=21 and 19) (**b**) or ZT12-14 (n=19 and 13) (**c**). **d** Representative traces of voltage-clamp recordings of *mWAKE*$^{+}$ SCN neurons of *mWAKE*$^{(Cre/+)}$*; Bmal1*$^{(+/+)}$ (grey) vs *mWAKE*$^{(Cre/+)}$*; Bmal1*$^{(-/-)}$ (yellow) at ZT12-14. Timing of GABA (1 mM) application is shown. **e**, GABA-evoked current in mWAKE^{SCN} neurons from *mWAKE*$^{(Cre/+)}$*; Bmal1*$^{(+/+)}$ (n=24) or *mWAKE*$^{(Cre/+)}$*; Bmal1*$^{(-/-)}$ (n=26) mice at ZT12-14. n represent individual cells from 4 animals.

Extended Data Figure 4 | Loss of mWAKE in SCN does not cause hyperactivity
a, Schematic of the genomic structure of the *mWAKE* locus and insertion of *loxP* sites flanking exons 4 and 5 in the *mWAKE*$^{(flox)}$ allele. **b** and **d**, Locomotor activity profile over 24 hrs for *mWAKE*$^{(flox/-)}$ (grey, n=5) and *Six3*$^{(Cre/+)}$*; mWAKE*$^{(flox/-)}$ (red, n=6) under

LD (**b**) or DD (**d**) conditions. **c** and **e**, Total locomotor activity from the mice in (**b**) at ZT0-12 and ZT12-24 or (**d**) at CT0-12 and CT12-24.

Extended Data Figure 5 | mWAKE$^+$ DMH neurons regulate sleep/wake states and receive neuromodulatory input

a, Schematic of whole-cell patch-clamp recordings from mWAKE$^+$ neurons (tdTomato$^+$ neurons in *mWAKE$^{(Cre/+)}$* or *mWAKE$^{(Cre/Cre)}$* mice), in the DMH region (top). Representative images of a biocytin- and tdTomato-labeled cell post-recording (bottom). Scale bar denotes 25 µm. **b** and **c**, *f-I* curves for mWAKE^{DMH} neurons from *mWAKE$^{(Cre/+)}$* vs *mWAKE$^{(Cre/Cre)}$* mice at ZT0-2 (n=11 and 11, Day) (**b**) or ZT12-14 (n=8 and 8, Night) (**c**). n represent individual cells from 4 animals. **d** and **e**, NREM (**d**) and REM sleep (**e**) (% per hr bin) for *mWAKE$^{(Cre/+)}$* mice with AAV-DIO-DREADD-Gq injected bilaterally into the DMH, following IP injection of vehicle (grey) or 1 mg/kg CNO (red) at ZT6. Note these are data are from the same mice as in Fig. 2m. **f**, Total locomotor activity (total number of break breaks on the X and Y axis) of *mWAKE$^{(Cre/+)}$* mice with bilateral injections of AAV-DIO-DREADD-Gi into the DMH in the 4 hrs following IP injection of vehicle vs CNO (1 mg/kg) (n=4) at ZT10. **g**, Representative short-time Fourier transform spectrograms of 4 hrs of recorded EEG activity, starting after IP injection of vehicle alone (above) or 3 mg/kg CNO (below) at ZT10, from *mWAKE$^{(Cre/+)}$* mice injected with AAV-DIO-DREADD-Gi bilaterally into the DMH. **h-j**, Wakefulness (**h**), NREM (**i**), and REM (**j**) amount for the mice described in (**g**), plotted as % time in 1 hr bins following IP injection of vehicle (grey) or 3 mg/kg CNO (red) (n=4) at ZT10.

Extended Data Figure 6 | Single cell sequencing

a, UMAP plot showing distribution of mWAKE$^+$ cells in individual scRNA-Seq libraries and distribution of mWAKE$^+$ neurons and ependymal cells. **b**, Violin plot showing distribution of number and mean (black dot) of genes (top) and total mRNAs (bottom, calculated by the number of unique molecular identifiers/UMIs) in individual scRNA-Seq libraries. **c**, Heatmap showing key marker genes that were used to identify spatial location of each mWake$^+$ neuronal cluster. **d-g**, Representative images of FISH experiments using RNAscope probes against *tdTomato*, *Vglut2*, and/or *Vgat* mRNA in

the DMH (**d** and **e**) and SCN (**f** and **g**) of *mWAKE*$^{(Cre/+)}$ mice. Nuclei counterstained with DAPI. Scale bars denote 200 μm in **d** and **f** and 20 μm in **e** and **g**. **h**, Heatmap showing normalized gene expression of specific norepinephrine (norepi), acetylcholine (Ach), dopamine (DA), orexin (Orx), and histamine (HA) receptors expressed in mWAKE^{+} neuronal clusters.

Extended Data Figure 7 | Additional data related to neuromodulatory input to mWAKE^{+} DMH neurons

a, Schematic showing dual injections of the Cre-dependent rabies helper virus (AAV-FLEX-G) and the rabies-ΔG-BFP (RabV-ΔG-BFP) virus into the DMH of *mWAKE*$^{(Cre/+)}$ mice. **b-e**, Representative images of BFP fluorescence and anti-histamine decarboxylase (HDC) (**b**), anti-Orexin A (OxA) (**c**), anti-choline acetyltransferase (CHAT) (**d**), or anti-tyrosine hydroxylase (TH) (**e**), as well as merged images from TMN (**b**), lateral hypothalamus (LH) (**c**), BF (**d**), or LC (**e**) regions, from the mice described in (**a**). Scale bar denotes 20 μm in (**b**), 50 μm in (**c**), 20 μm in (**d**), 20 μm in (**e**). **f**, Schematic showing patch-clamp recordings of mWAKE^{DMH} neurons from *mWAKE*$^{(Cre/+)}$ mice, following application of histamine. **g**, Representative membrane potential traces from the recordings described in (**f**) in the presence of synaptic blockers at ZT12-14. **h**, Spontaneous mean firing rate of mWAKE^{DMH} neurons following application of 20 μM histamine (n=8 and 9 cells) at ZT12-14 in the absence ("- synaptic blockers") or presence ("+ synaptic blockers") of 20 μM CNQX, 50 μM AP5, and 10 μM picrotoxin. The neurons recorded from in **(h)** were derived from 3 animals.

Extended Data Figure 8 | The mWAKE network is critical for arousal

a-e, mWAKE^{+} cells (identified by tdTomato fluorescence in *mWAKE*$^{(Cre/+)}$ mice) are near but distinct from cholinergic neurons in the BF (**a**), orexinergic neurons in the LH (**b**), histaminergic neurons in the TMN (**c**), serotonergic neurons in the DR (**d**), and noradrenergic neurons in the LC (**e**), defined by anti-CHAT, anti-Orx, anti-HDC, anti-TPH2, and anti-TH labeling, respectively. Scale bars denote 200 μm (**a**), 100 μm (**b**), 200 μm (**c**), 100 μm (**d**), and 100 μm (**e**). **f-i**, Representative images of FISH experiments using RNAscope probes against *tdTomato*, *Vglut2*, and/or *Vgat* mRNA in

the LC (**f** and **g**) and BF (**h** and **i**) of *mWAKE*$^{(Cre/+)}$ mice. Nuclei counterstained with DAPI. Scale bars denote 50 μm. **j**, Representative images of anti-HA labeling of cells and processes expressing LSL-DREADD-Gi, tdTomato fluorescence from mWAKE^{+} cells, or merged images in representative brain regions from *mWAKE*$^{(Cre/+)}$ > *LSL-DREADD-Gi* mice. Sections containing suprachiasmatic nucleus (SCN), basal forebrain (BF), and cortex (Cx) are shown. Scale bar denotes 100 μm in (SCN), 20 μm in BF, and 100 μm in (Cx). **k**, Jitter plot showing behavioural classification of arousal level for *HDC*$^{(Cre/+)}$*; LSL-Gi* mice following injection of vehicle or 1 mg/kg CNO (n=4). **l**, Representative short-time Fourier transform spectrograms of 12 hrs of recorded EEG activity from *HDC*$^{(Cre/+)}$; *LSL-Gi* mice after IP injection at ZT10 of vehicle alone (above) or 1 mg/kg CNO (below). **m**, Short-time Fourier transform spectrograms for a *HDC*$^{(Cre/+)}$*;LSL-Gi* mouse starting 150 mins after injection of vehicle (left) or 1 mg/kg CNO (right) at ZT10. The red box indicates the time window used for spectral and amplitude analysis in (**n**) and (**o**). **n**, Welch's power spectral density estimates as a percentage of total EEG power across 10 min, averaged across multiple EEG traces. *HDC*$^{(Cre/+)}$*;LSL-Gi* injected with vehicle (grey) or 1 mg/kg CNO (green), n=3. Inset shows a plot of delta-band power as a percentage of total EEG power. **o**, EEG trace amplitude (plotted as normalized root mean square (RMS)) for the animals described in (**n**).

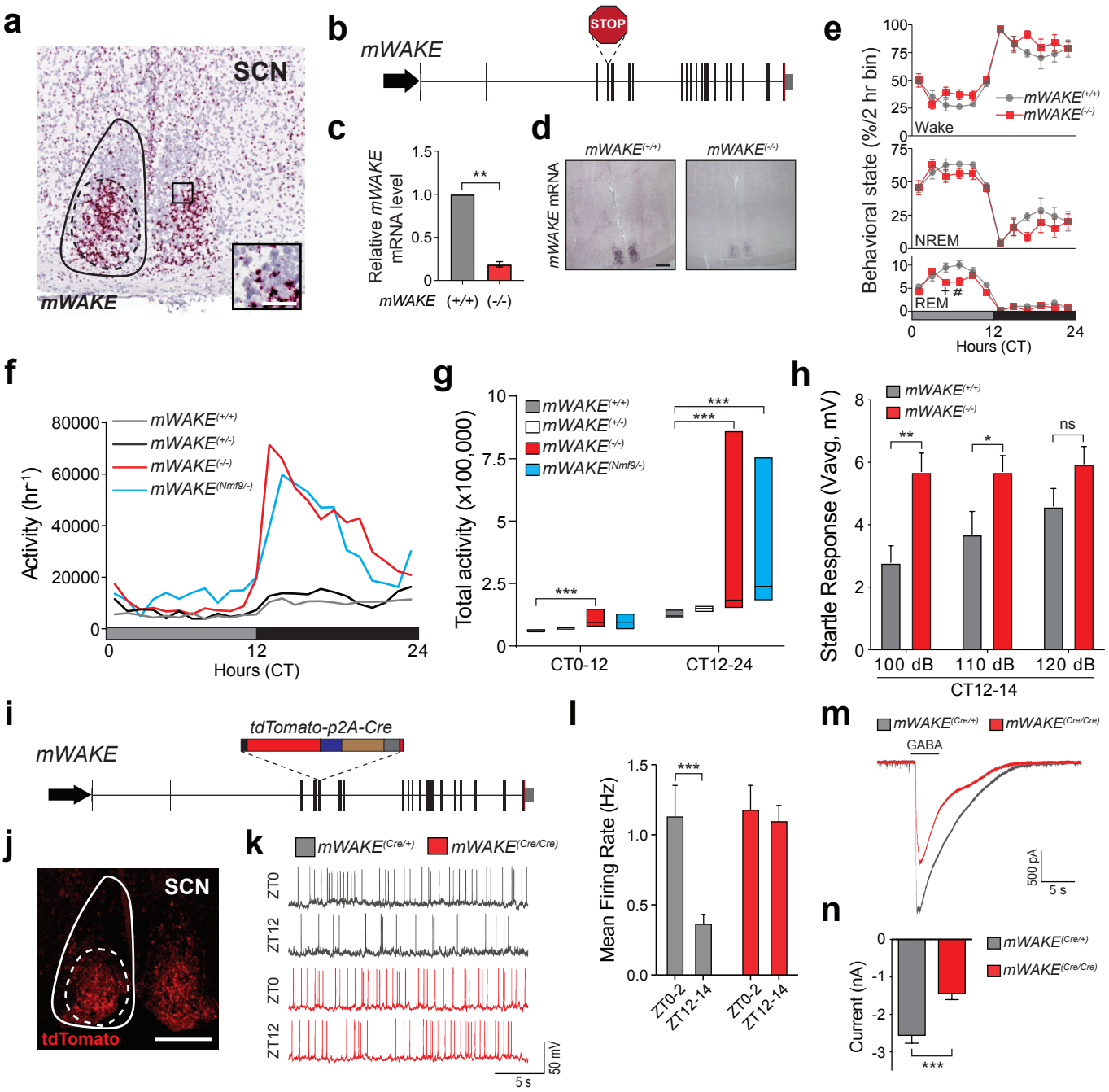
a
SCN
mWAKE
b
mWAKE
STOP
c
Relative mWAKE mRNA level
1.5
1.0
0.5
0
**
mWAKE (+/+) (-/-)
d
mWAKE(+/+)
mWAKE(-/-)
mWAKE mRNA
e
Behavioral state (%/2 hr bin)
100
75
50
25
0
Wake
mWAKE(+/+)
mWAKE(-/-)
NREM
10
5
REM
+ #
0
12
24
Hours (CT)
f
Activity (hr-1)
80000
60000
40000
20000
0
mWAKE(+/+)
mWAKE(+/-)
mWAKE(-/-)
mWAKE(Nmf9/-)
0
12
24
Hours (CT)
g
Total activity (x100,000)
10
7.5
5
2.5
0
mWAKE(+/+)
mWAKE(+/-)
mWAKE(-/-)
mWAKE(Nmf9/-)
***
***
***
CT0-12
CT12-24
h
Startle Response (Vavg, mV)
8
6
4
2
0
mWAKE(+/+)
mWAKE(-/-)
**
*
ns
100 dB
110 dB
120 dB
CT12-14
i
mWAKE
tdTomato-p2A-Cre
j
SCN
tdTomato
k
mWAKE(Cre/+)
mWAKE(Cre/Cre)
ZT0
ZT12
ZT0
ZT12
50 mV
5 s
l
Mean Firing Rate (Hz)
1.5
1.0
0.5
0.0
***
ZT0-2
ZT12-14
ZT0-2
ZT12-14
m
mWAKE(Cre/+)
mWAKE(Cre/Cre)
GABA
500 pA
5 s
n
Current (nA)
0
-1
-2
-3
mWAKE(Cre/+)
mWAKE(Cre/Cre)
***

Figure 1

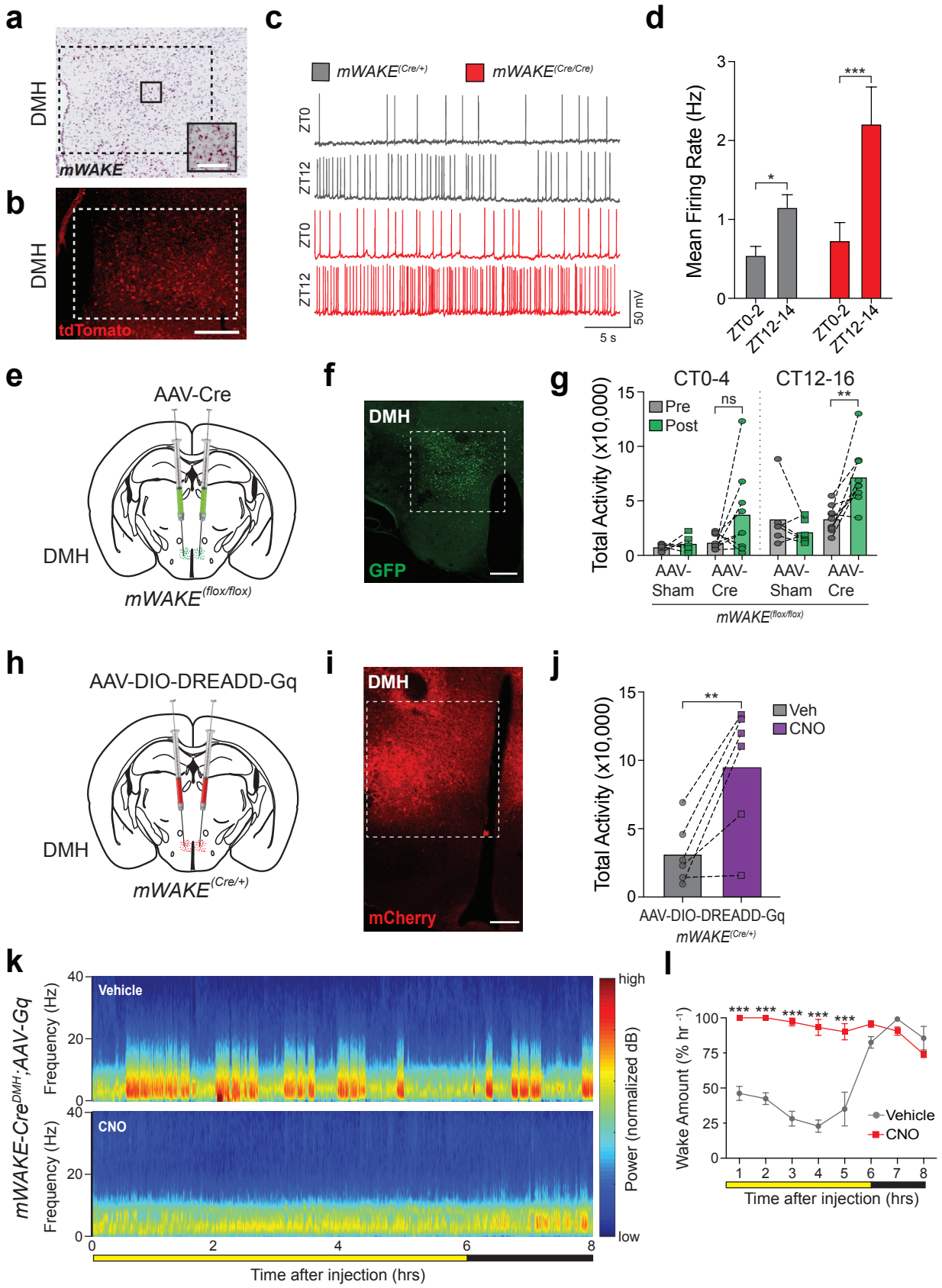
a
DMH
mWAKE
b
DMH
tdTomato
c
mWAKE(Cre/+)
mWAKE(Cre/Cre)
ZT0
ZT12
ZT0
ZT12
50 mV
5 s
d
Mean Firing Rate (Hz)
***
*
ZT0-2
ZT12-14
ZT0-2
ZT12-14
e
AAV-Cre
DMH
mWAKE(flox/flox)
f
DMH
GFP
g
CT0-4
CT12-16
Pre
Post
ns
**
Total Activity (x10,000)
AAV-Sham
AAV-Cre
AAV-Sham
AAV-Cre
mWAKE(flox/flox)
h
AAV-DIO-DREADD-Gq
DMH
mWAKE(Cre/+)
i
DMH
mCherry
j
**
Veh
CNO
Total Activity (x10,000)
AAV-DIO-DREADD-Gq
mWAKE(Cre/+)
k
mWAKE-CreDMH;AAV-Gq
Vehicle
CNO
Frequency (Hz)
Power (normalized dB)
high
low
Time after injection (hrs)
l
***
Wake Amount (% hr -1)
Vehicle
CNO
Time after injection (hrs)

Figure 2

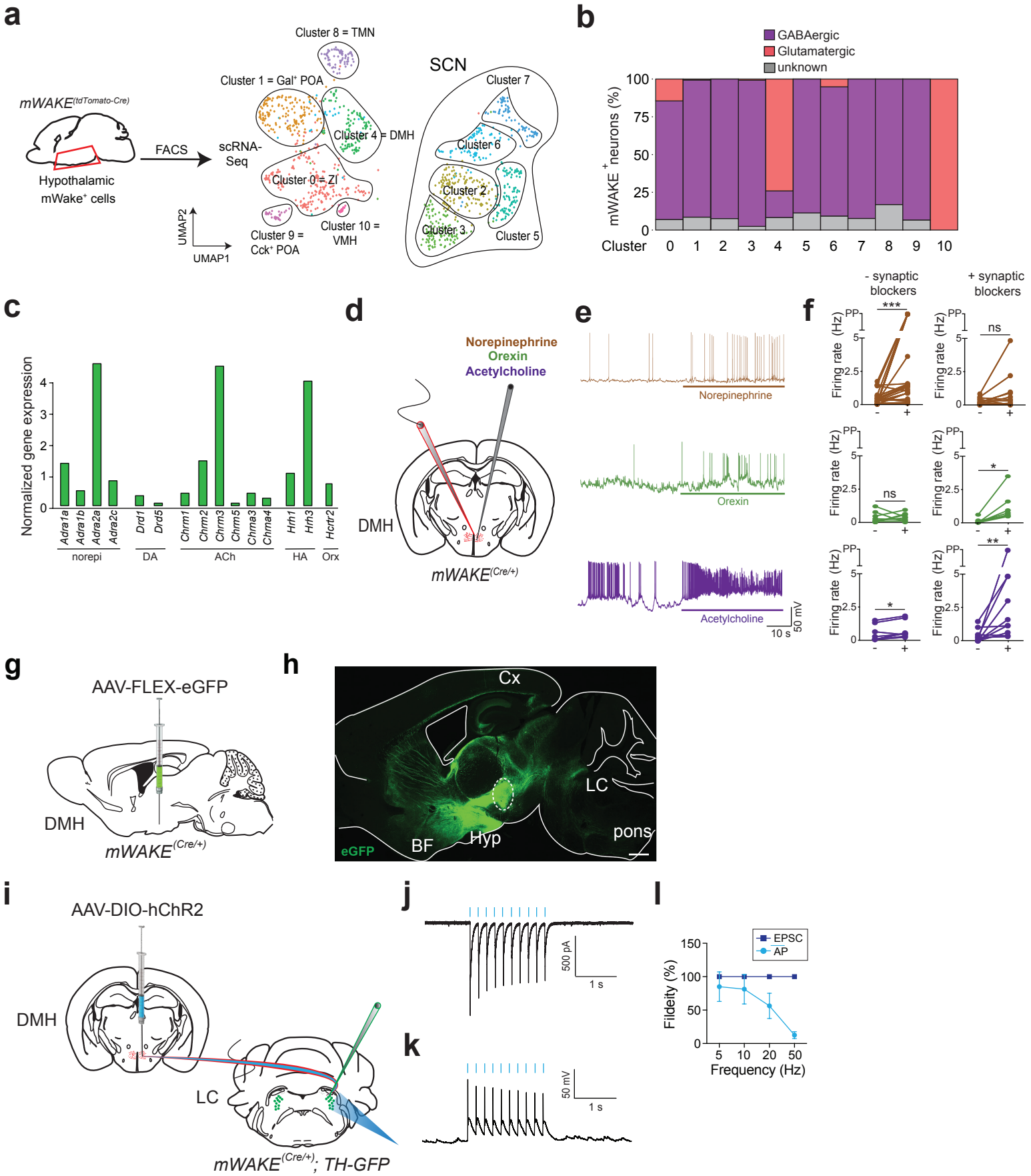
a
mWAKE(tdTomato-Cre)
Hypothalamic mWake+ cells
FACS
scRNA-Seq
Cluster 8 = TMN
Cluster 1 = Gal+ POA
Cluster 4 = DMH
Cluster 0 = ZI
Cluster 9 = Cck+ POA
Cluster 10 = VMH
UMAP2
UMAP1
SCN
Cluster 7
Cluster 6
Cluster 2
Cluster 3
Cluster 5
b
GABAergic
Glutamatergic
unknown
mWAKE+ neurons (%)
Cluster
c
Normalized gene expression
Adra1a
Adra1b
Adra2a
Adra2c
Drd1
Drd5
Chrm1
Chrm2
Chrm3
Chrm5
Chrma3
Chrma4
Hrh1
Hrh3
Hcrtr2
norepi
DA
ACh
HA
Orx
d
Norepinephrine
Orexin
Acetylcholine
DMH
mWAKE(Cre/+)
e
Norepinephrine
Orexin
Acetylcholine
50 mV
10 s
f
- synaptic blockers
+ synaptic blockers
Firing rate (Hz)
PP
***
ns
*
**
g
AAV-FLEX-eGFP
DMH
mWAKE(Cre/+)
h
Cx
LC
pons
BF
Hyp
eGFP
i
AAV-DIO-hChR2
DMH
LC
mWAKE(Cre/+); TH-GFP
j
500 pA
1 s
k
50 mV
1 s
l
EPSC
AP
Fidelity (%)
Frequency (Hz)

Figure 3

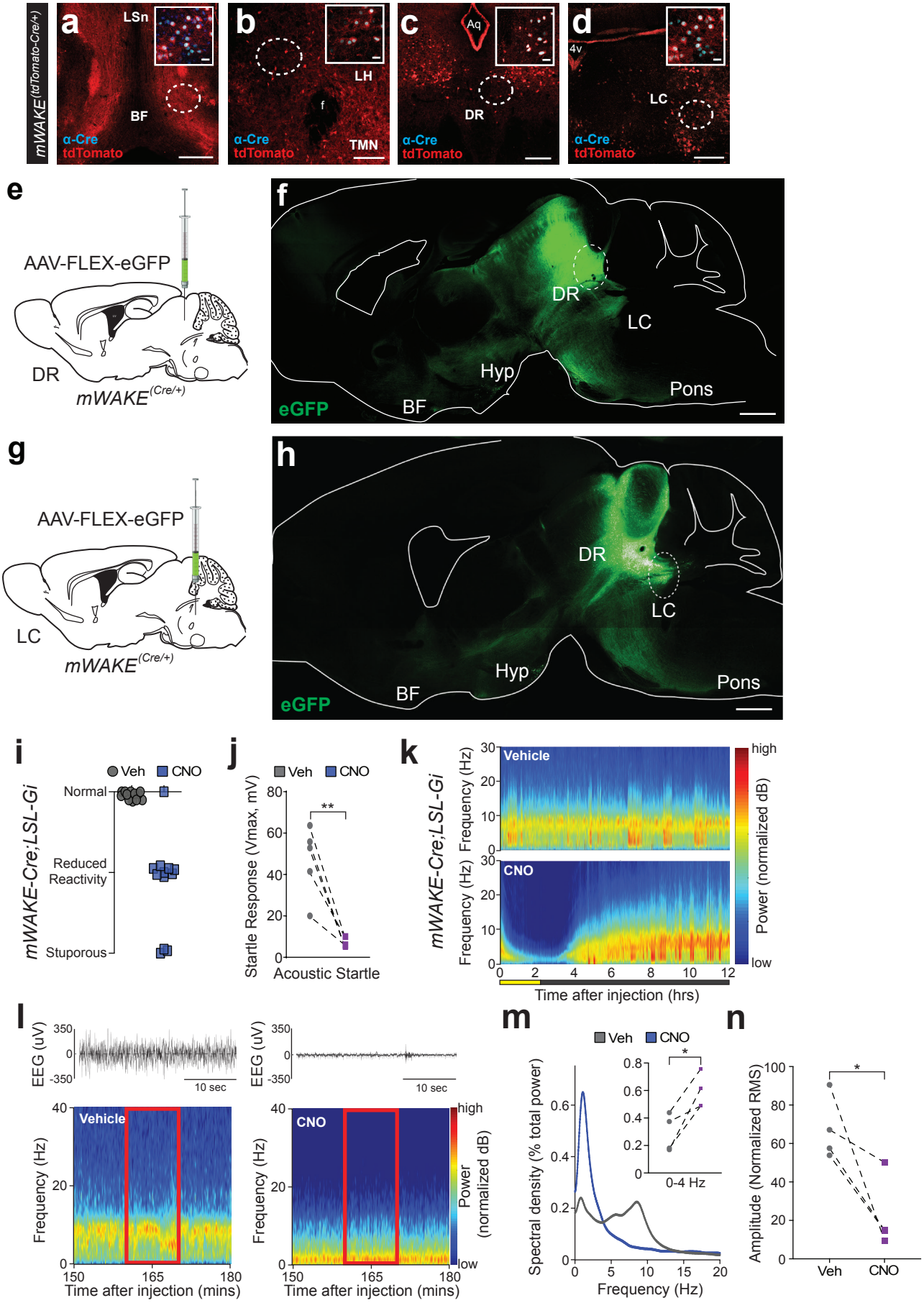
mWAKE(tdTomato-Cre/+)
a
LSn
BF
α-Cre
tdTomato
b
LH
f
TMN
c
Aq
DR
d
4v
LC
e
AAV-FLEX-eGFP
DR
mWAKE(Cre/+)
f
DR
LC
Hyp
Pons
BF
eGFP
g
AAV-FLEX-eGFP
LC
mWAKE(Cre/+)
h
DR
LC
Hyp
Pons
BF
eGFP
i
Veh
CNO
Normal
Reduced Reactivity
Stuporous
mWAKE-Cre;LSL-Gi
j
Veh
CNO
Startle Response (Vmax, mV)
**
Acoustic Startle
k
mWAKE-Cre;LSL-Gi
Vehicle
CNO
Frequency (Hz)
Power (normalized dB)
high
low
Time after injection (hrs)
l
EEG (uV)
10 sec
Vehicle
CNO
Frequency (Hz)
Power (normalized dB)
Time after injection (mins)
m
Veh
CNO
Spectral density (% total power)
0-4 Hz
Frequency (Hz)
n
Amplitude (Normalized RMS)
Veh
CNO

Figure 4