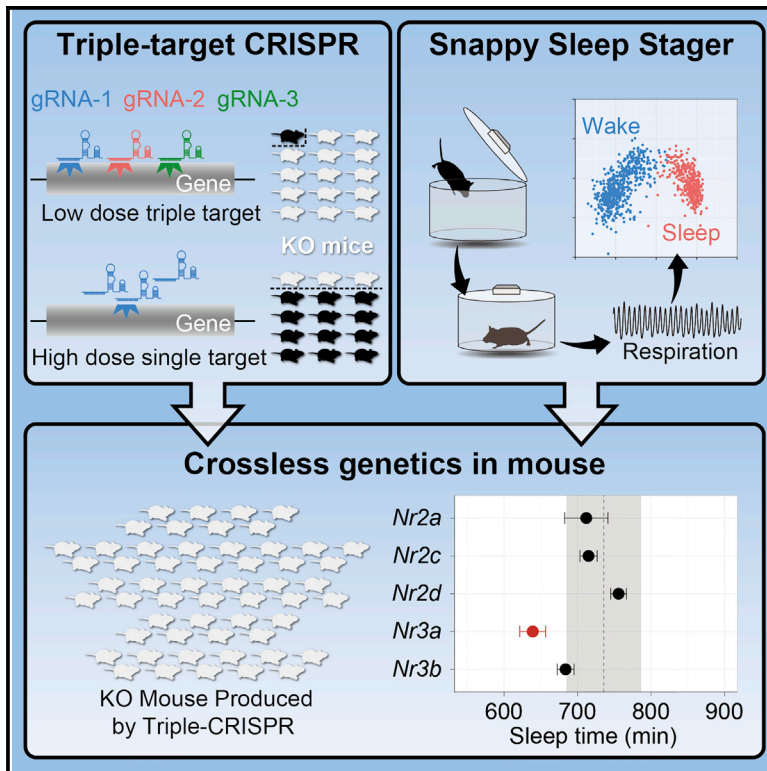


## Mammalian Reverse Genetics without Crossing Reveals *Nr3a* as a Short-Sleeper Gene

### Graphical Abstract



### Authors

Genshiro A. Sunagawa, Kenta Sumiyama, Maki Ukai-Tadenuma, Dimitri Perrin, ..., Koji L. Ode, Shigehiro Kuraku, Hiroki R. Ueda

### Correspondence

uedah-ky@umin.ac.jp

### In Brief

Sunagawa et al. present an application of mammalian reverse genetics without crossing by developing two methods. The authors improve KO mice production using a triple-target CRISPR and combine this with a non-invasive respiration-based automated sleep phenotyping system, the Snappy Sleep Stager. Combining these methods, the authors found that the *Nr3a* KO is a short-sleeper.

### Highlights

- A triple-target CRISPR method achieved almost perfect knockout efficiency
- SSS achieved non-invasive fully automated high-performance sleep phenotyping
- Genetics without crossing revealed *Nr3a* as a short-sleeper gene



# Mammalian Reverse Genetics without Crossing Reveals *Nr3a* as a Short-Sleeper Gene

Genshiro A. Sunagawa,<sup>1,8</sup> Kenta Sumiyama,<sup>2,8</sup> Maki Ukai-Tadenuma,<sup>1,8</sup> Dimitri Perrin,<sup>1,6,8</sup> Hiroshi Fujishima,<sup>1</sup> Hideki Ukai,<sup>1</sup> Osamu Nishimura,<sup>5</sup> Shoi Shi,<sup>3</sup> Rei-ichiro Ohno,<sup>3</sup> Ryohei Narumi,<sup>1</sup> Yoshihiro Shimizu,<sup>7</sup> Daisuke Tone,<sup>3</sup> Koji L. Ode,<sup>3</sup> Shigehiro Kuraku,<sup>5</sup> and Hiroki R. Ueda<sup>1,3,4,\*</sup>

<sup>1</sup>Laboratory for Synthetic Biology, RIKEN Quantitative Biology Center, 1-3 Yamadaoka, Suita, Osaka 565-0871, Japan

<sup>2</sup>Laboratory for Mouse Genetic Engineering, RIKEN Quantitative Biology Center, 1-3 Yamadaoka, Suita, Osaka 565-0871, Japan

<sup>3</sup>Department of Systems Pharmacology, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

<sup>4</sup>CREST, Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan

<sup>5</sup>Phyloinformatics Unit, RIKEN Center for Life Science Technologies, 2-2-3 Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan

<sup>6</sup>School of Electrical Engineering and Computer Science, Science and Engineering Faculty, Queensland University of Technology, GPO Box 2434, Brisbane, QLD 4001, Australia

<sup>7</sup>Laboratory for Cell-Free Protein Synthesis, RIKEN Quantitative Biology Center, 6-2-3, Furuedai, Suita, Osaka 565-0874, Japan

<sup>8</sup>Co-first author

\*Correspondence: [uedah-ky@umin.ac.jp](mailto:uedah-ky@umin.ac.jp)

<http://dx.doi.org/10.1016/j.celrep.2015.12.052>

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## SUMMARY

The identification of molecular networks at the system level in mammals is accelerated by next-generation mammalian genetics without crossing, which requires both the efficient production of whole-body biallelic knockout (KO) mice in a single generation and high-performance phenotype analyses. Here, we show that the triple targeting of a single gene using the CRISPR/Cas9 system achieves almost perfect KO efficiency (96%–100%). In addition, we developed a respiration-based fully automated non-invasive sleep phenotyping system, the Snappy Sleep Stager (SSS), for high-performance (95.3% accuracy) sleep/wake staging. Using the triple-target CRISPR and SSS in tandem, we reliably obtained sleep/wake phenotypes, even in double-KO mice. By using this system to comprehensively analyze all of the *N*-methyl-*D*-aspartate (NMDA) receptor family members, we found *Nr3a* as a short-sleeper gene, which is verified by an independent set of triple-target CRISPR. These results demonstrate the application of mammalian reverse genetics without crossing to organism-level systems biology in sleep research.

## INTRODUCTION

The system-level identification of molecular networks in organisms is an important challenge in biology (Kitano, 2002a, 2002b). Classical reverse genetics requires several generations of animal crosses to produce mutant animals of sufficient quality and quantity for phenotype analysis. Conventional methods for producing knockout (KO) mice usually involve targeting-vector

construction (~2 weeks), the introduction of target mutations into embryonic stem cells (ESCs) by homologous recombination (a few weeks), and the injection of the mutant ESCs into wild-type blastocysts to produce chimera mice (~3 weeks). If the mutant ESCs contribute to the germline of the newborn chimera mice, their next-generation offspring will possess a heterozygous mutation (~3 months). Further crossings of the offspring (several months to years; at least 3 months per generation) will produce mice with completely homozygous KO mutations on an inbred genomic background, which is required for reliable phenotype analysis. Thus, conventional methods require substantial amounts of time, space, and effort to knock out even a single gene (Nagy, 2003). Therefore, identification of molecular networks comprehensively in organisms will require “next-generation” genetics, i.e., genetic alterations without crossing. Toward this goal, both highly efficient (>90%) mutant production in a single generation and highly accurate (>90%) phenotype analyses of mutant animals are needed.

The efficient production of biallelic KO mice can be facilitated by recently developed genome editing techniques, including ZFNs (zinc-finger nucleases) (Carbery et al., 2010), TALENs (transcription-activator-like effector nucleases) (Sung et al., 2013), and the CRISPR/Cas nuclease systems (Hsu et al., 2014; Wang et al., 2013). Unlike the protein-based ZFNs and TALENs, the CRISPR/Cas system uses RNA-based DNA recognition derived from a bacterial adaptive immune system (Horvath and Barrangou, 2010; Wiedenheft et al., 2012). This method accelerates the generation of KO animals via the co-injection of RNA encoding the Cas9 protein and target-locus-specific guide RNAs (gRNAs) into embryos (Fujii et al., 2013; Wang et al., 2013). Several modifications of the CRISPR/Cas9 system have been also introduced to improve the efficiency and specificity of targeted mutations in a genome (Fu et al., 2014; Ran et al., 2013; Zhou et al., 2014). However, two problems remain: (1) first-generation mice often contain a mosaic of wild-type and KO cells, and

(2) the rate of complete biallelic mutant mice generated is relatively low (usually ~60%–80% at best). Therefore, the highly efficient (>90%) production of whole-body biallelic KO in a single generation remains a fundamental challenge for organism-level systems biology.

The comprehensive identification of molecular circuits at the organism level also requires accurate (>90%) phenotype analysis. In neuroscience, sleep/wake behavior is an intriguing phenotype, because sleep disorders (e.g., insomnia or hypersomnia) are sensitive and informative symptoms of almost all psychological disorders. Sleep/wake states have been characterized in humans by electroencephalography (EEG) and electromyography (EMG). For example, during sleep, EEG mostly displays high-amplitude, low-frequency fluctuations, whereas during waking, it exhibits low-amplitude, high-frequency fluctuations. On the other hand, EMG displays low-amplitude fluctuations during sleep and high-amplitude fluctuations during waking. These characteristic EEG/EMG patterns during sleep and waking are preserved in mammals and can be measured by electrodes surgically implanted in the brain and muscles. However, such recording requires special surgical skills, and the surgery is highly invasive, requiring a long recovery period (more than 10 days) after implantation before sleep/wake recording. Furthermore, the EEG/EMG data are often manually annotated and classified into sleep/wake phenotypes by visual assessment, which can be time consuming and somewhat subjective. Therefore, sleep/wake phenotyping has been a low-throughput method; for comprehensive studies, a scalable, non-invasive, fully automated sleep/wake recording method is needed.

In this study, we developed a simple theory to predict the minimum efficacy of different CRISPR methods for the highly efficient production of whole-body biallelic KO mice. As we predicted, triple-target CRISPR elicited almost perfect (~96%–100%) whole-body KO of the *Tyrosinase* (*Tyr*) gene, which is functionally evaluated by animal coat color. This KO efficiency was confirmed using three independent sets of gRNAs. The highly efficient production of whole-body KO by the triple-target CRISPR method also enabled us to obtain clock mutant phenotypes reliably, not only of *Bmal1* single-KO mice but also of *Cry1/Cry2* or *Per1/Per2* double-KO mice. For accurate phenotype analysis, we developed a respiration-based, non-invasive, fully automated system, the Snappy Sleep Stager (SSS), which enabled the high-performance analysis (95.3% accuracy) of sleep/wake phenotypes. Using the triple-target CRISPR and SSS methods, we analyzed the *N*-methyl-*D*-aspartate (NMDA) receptor (NMDAR) family comprehensively and found that the *Nr3a* KO mouse is a short sleeper (96.7 min/day less than wild-type). Finally, we developed a freely accessible web-based database that contains the triple-target candidates for 81.2% of the genes in the mouse genome.

## RESULTS

### Triple-Target CRISPR Achieves Almost Perfect KO Efficiency

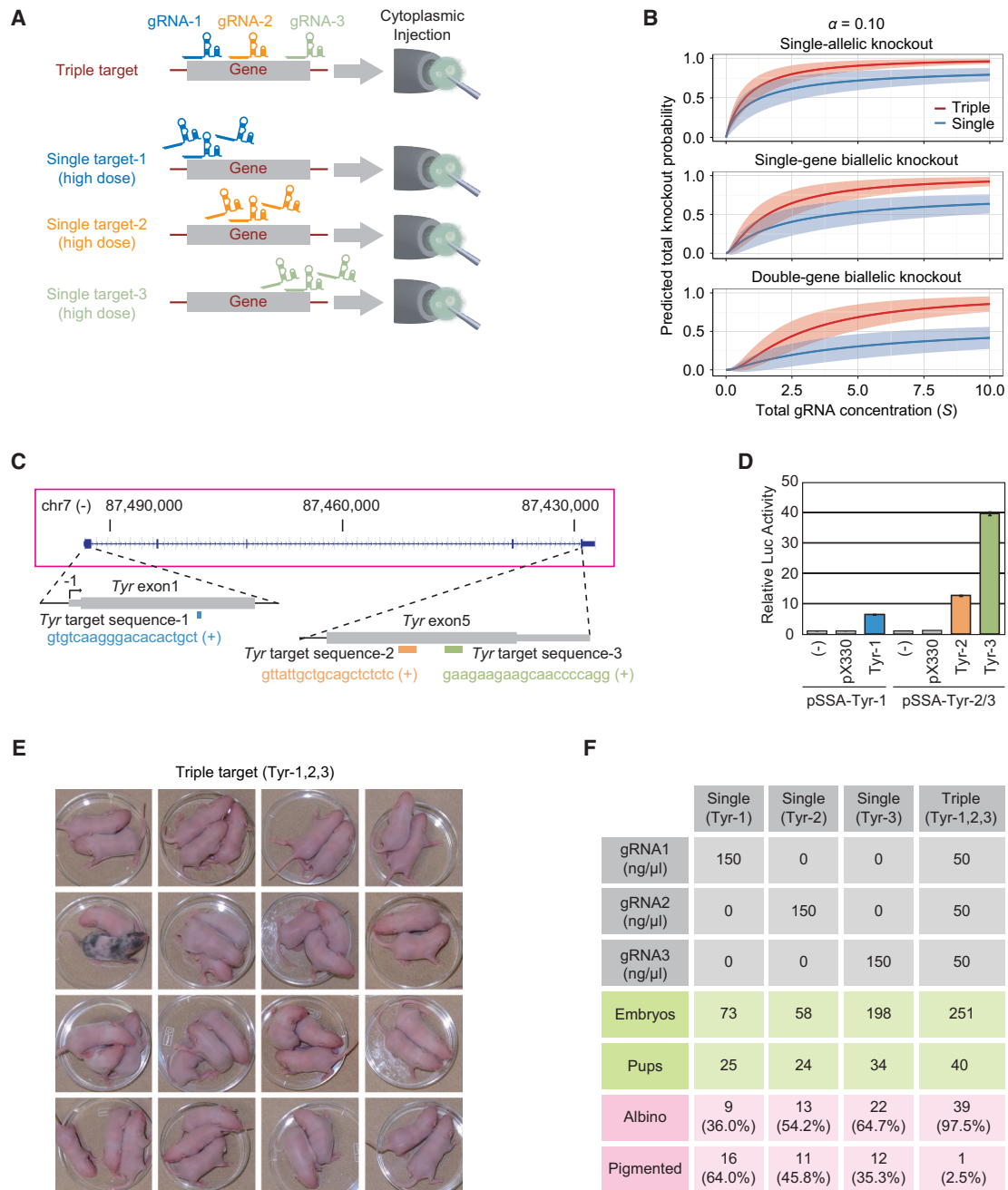
To obtain a highly efficient method for producing biallelic KO mice, we first constructed a simple computational model to es-

timate the minimum efficacy of different CRISPR methods (Supplemental Experimental Procedures). According to this computational model, a multiple-target CRISPR strategy, in which multiple gRNAs target the same gene, is more efficient than the high-concentration CRISPR strategy, in which a multiple-fold higher concentration of a single gRNA targets the gene of interest (Figures 1A and 1B; Figures S1A and S1B). This finding is consistent with a previous report on dual-target CRISPR (Zhou et al., 2014). Our model further predicted that triple-target CRISPR would achieve 79%–97% efficiency for the biallelic KO of a single gene (Figure S1C). Importantly, if it achieves more than 90% efficiency for the biallelic KO of a single gene, it can also achieve more than 80% efficiency even for the biallelic KO of dual genes, which is sufficient for reliable analysis of behavioral phenotypes.

To test the model's predictions, we selected *Tyr* as a target gene and C57BL/6N as a basal inbred strain, because if this gene is knocked out biallelically, the black coat color of C57BL/6N becomes white. We designed three gRNAs for three target sequences of the *Tyr* gene (gRNA1, gRNA2, and gRNA3 for targets Tyr-1, Tyr-2, and Tyr-3, respectively; Figures 1C and S1D). We confirmed that the designed gRNAs had high cleavage efficiency (~7-fold to ~40-fold compared to control) at least in cellulo by using a single strand annealing (SSA) assay (Figures 1D and S1E), in which the DNA cleavage of a target sequence by the CRISPR/Cas9 system induced the recombination of incomplete fragments of firefly *Luciferase* (*Luc*), resulting in an entire *Luc* gene and enhanced bioluminescence.

We then estimated two parameters of the model, an effective total gRNA concentration compared to the dissociation constant of gRNA ( $S$ ) and a recovery rate of a mutation ( $\alpha$ ). When injected into fertilized eggs, the single-target strategy using each of the three gRNAs alone at 150 ng/ $\mu$ l exhibited only moderate efficacy (36.0%, 54.2%, and 64.7%, for gRNA1, gRNA2, and gRNA3, respectively; Figures 1F and S1G). Based on these results,  $S$  and  $\alpha$  were estimated to be 5.72 and 0.136, respectively. This allowed us to predict the minimum efficacy of the triple-target CRISPR method for whole-body biallelic KO mice as 82.6% (Figure S1F). In fact, the triple-target strategy using the mixture of Cas9 mRNA (100 ng/ $\mu$ l) and three gRNAs (gRNA1, gRNA2, and gRNA3, each at 50 ng/ $\mu$ l) achieved almost perfect efficiency (97.5%; Figures 1E and 1F). We also performed qPCR of the target region including the CRISPR-target sequence and confirmed that five out of six CRISPR-target regions from two *Tyr* KO mice were not detectable. These qPCR results matched with each other among the genomic DNAs extracted from brains, scalps, and tails (Figure S1H), which is consistent with the observed whole-body biallelic KO phenotype (i.e., white coat color) of the *Tyr* gene. We noted that observed KO efficacy (97.5%) is more than the predicted minimum efficiency (82.6%) (Figure S1F). We speculated that the increased KO efficacy would be because of irreversible mutations such as deletion between multiple CRISPR target sites.

To investigate its genomic effect, we performed exome sequencing; results were that 91.6%–95.6% of target bases were covered more than ten times (Table S5). As a result, it demonstrated on-target deletions of genomic stretches with variable lengths (Figure 2) and no off-target mutation. Thus, the



**Figure 1. A Triple-Target CRISPR Method for Highly Efficient Production of Whole-Body Biallelic KO Mice**

(A) Schematic diagram of triple-target and single-target CRISPR methods. Each method uses the same amount of total gRNA.

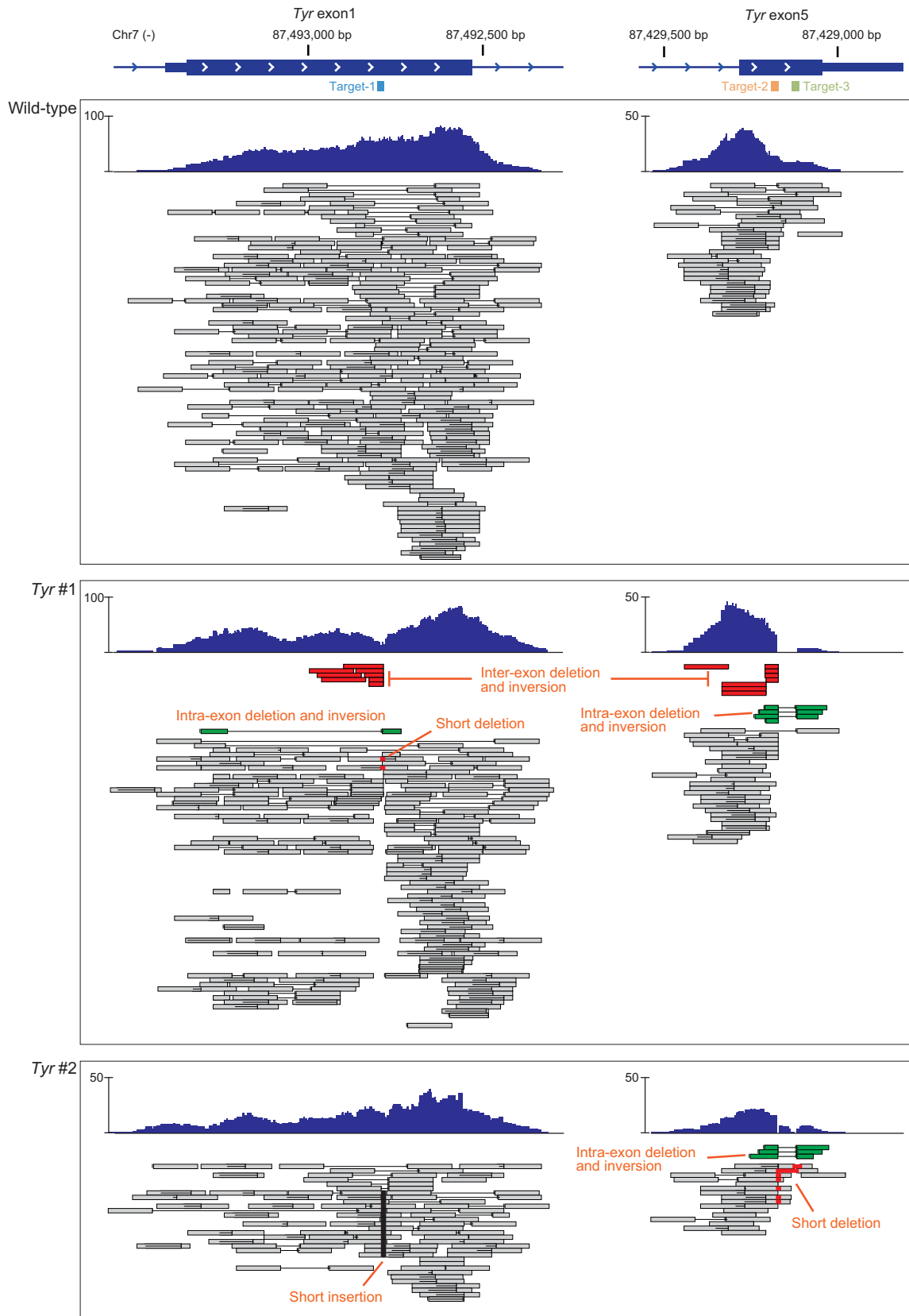
(B) Computational simulation predicting how efficiently one or more alleles are cut out when different numbers of gRNAs are used against the same target gene. The triple-target (red line) and single-target (blue line) methods were compared. The line and shaded area around the line indicate the mean and the SD of 1,000 simulations, respectively. The difference in efficiency between triple-target and single-target strategies became more apparent as the number of target alleles increased.

(C) Target sequences of the gRNAs for KO of the *Tyrosinase* (*Tyr*) gene. The short colored bars (blue, orange, and green) indicate the 20-base target sequences. The target sequences were on the sense strand of the genomic DNA (plus sign refers to the gene's sense strand).

(D) SSA assay for the gRNAs of the *Tyr* gene. Relative luciferase activities from the SSA vectors (pSSA-Tyr-1 and pSSA-Tyr-2) were measured. The SSA vector was transfected into 293T cells with empty vector (-), pX330 Cas9 vector without any gRNA, or pX330 Cas9 and *Tyr* gRNA (Tyr-1, Tyr-2, or Tyr-3). The relative Luciferase activity for each sample was scaled so that the activity with empty vector (-) was defined as 1. Error bars represent SD (n = 3).

(E) *Tyr* KO mice created by the triple-target CRISPR method. The coat color showed the biallelically knocked out mice.

(F) Comparison of the single- and triple-target CRISPR methods. The table shows the gRNA injection condition and resulting phenotypes (Figure S1G). Embryos, number of injected and transferred embryos; Pups, number of pups born (total); Albino, pups with albino coat color (biallelic knockout); Pigmented, pups with mosaic or wild-type coat color.



(legend on next page)

triple-target CRISPR method could produce biallelic KO mice with >90% efficiency in a single generation. Furthermore, undetectable expression of TYR protein was confirmed by mass spectrometry (MS) comparing with another independently produced triple-target CRISPR KO strain. TYR protein was not detectable in two *Tyr* KO mice, whereas it was detected in one *Nr3a* KO mouse (Figure S1I). Importantly, since a triple-target CRISPR method achieves high efficiency (>90%) for the biallelic KO of a single gene, it has a potential to achieve efficiency higher than 80% even for the biallelic KO of dual genes, which is sufficient for reliable analysis of behavioral phenotypes.

### The SSS Enables Non-invasive, Fully Automated, High-Performance Sleep/Wake Phenotyping

To identify system-level molecular circuits in organisms, both the efficient production of KO mice and simple but accurate phenotyping are necessary. The standard method for phenotyping mammals' sleep/wake states is to record EEG/EMG data. Each short-period segment of EEG/EMG data, called an "epoch," is then annotated into one of three sleep/wake states ("REM [rapid eye movement] sleep," "NREM [non-REM] sleep," and "awake"), determined by the EEG/EMG pattern (Figure S2A). Here, to obtain a non-invasive and fully automated method for simple and accurate sleep/wake phenotype analysis, we developed the SSS (Figure 3A). For this method, we first focused on the non-invasive phenotyping of sleep/wake states by recording the respiration of an animal, since the different states exhibit unique respiration patterns (Figure S2A).

Among the current methods for non-invasively recording the respiration of mice (Glaab et al., 2007; Gordon and Ali, 1984), whole-body plethysmography (WBP) is one of the most popular and promising. The correlation between respiration patterns and EEG/EMG-based sleep/wake states was previously analyzed (Hernandez et al., 2012). However, previous studies mainly examined short recording periods (less than 1 day) during different states of consciousness, so the volume of the WBP chamber could be relatively small. Although a smaller chamber size increases the sensitivity of the respiration recording, too small a chamber decreases the animal's comfort and can affect sleep/wake behavior. Therefore, no system has been developed for recording for a week or more, which is critical for accurate sleep/wake phenotyping. Increasing the size of the chamber requires a compensatory increase in the signal/noise ratio of the WBP system itself.

To optimize the WBP system, we modeled it as an equivalent electrical circuit (Figure 3B; Supplemental Experimental Procedures for the detailed equations)—essentially a band-pass filter for a certain frequency of pressure change. By computer simulation, we optimized the resistance and capacitance in the WBP system to band-pass the 1- to 10-Hz flow (Figure 3C), which is the physiological range for mouse respiration frequency. Based on this analysis, we constructed a high-performance WBP sys-

tem (Figure 3A), which exhibited a high signal/noise ratio for the 1- to 10-Hz flow and allowed us to record detailed, long-term respiration patterns of mice non-invasively.

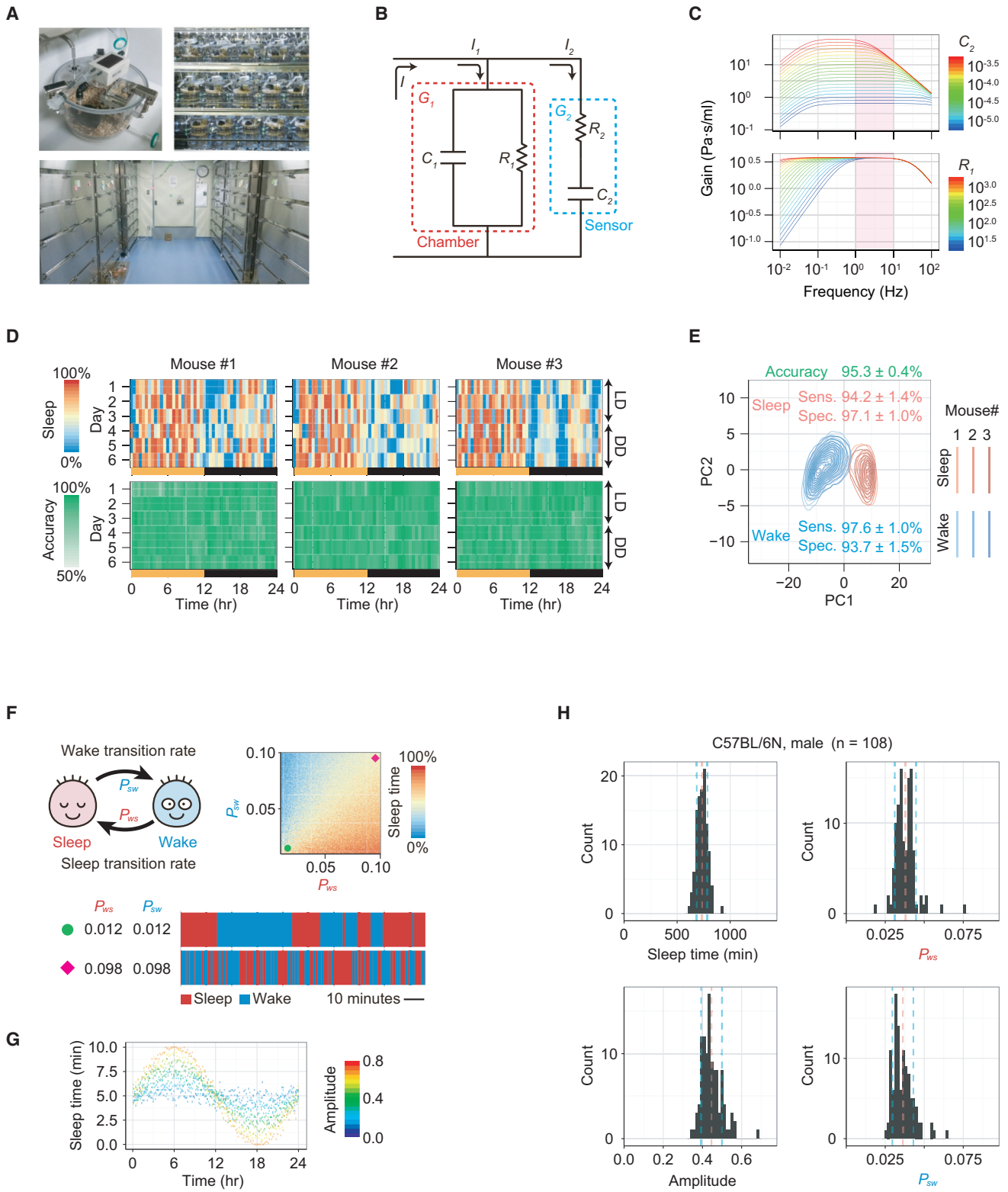
For simple and accurate sleep/wake phenotype analyses, fully automated annotation of the recorded data is also critical. We previously developed FASTER (an unsupervised fully automated sleep staging method for mice based on EEG/EMG recordings) (Sunagawa et al., 2013), which avoids the subjective manual annotation of EEG/EMG data. We applied the FASTER method to the recorded respiration data to create a fully automated annotation pipeline (Figure S2B; Supplemental Experimental Procedures). Each epoch is annotated as either "sleep" or "awake" fully automatically. Since REM state represented only ~5% of the total epochs, which agrees with past documentation (Chemelli et al., 1999), we focused on a two-state classification between sleep and awake states as the first-order approximation in this study. Simultaneous recordings of the respiration flow and EEG/EMG enabled us to estimate the performance of the SSS system (Figure 3D), resulting in the high accuracy of  $95.3 \pm 0.4\%$  ( $n = 3$ ) for the respiration-based sleep/wake phenotyping (Figures 3E and S2C). Importantly, the non-invasiveness of the recording and fully automated data analysis minimized the variability and subjectivity of the procedures, thereby making the sleep/wake phenotyping highly reproducible, as indicated by the overlapping probability densities of sleep/wake states among three individuals (Figure 3E). We also note that even "quiet awake," which was defined in this study as the awake state accompanied by the lower 33.3% of EMG values, was annotated accurately ( $95.3 \pm 1.3\%$ ;  $n = 3$ ) by the SSS system (Figure S2D). Therefore, the SSS provides a non-invasive, fully automated method for sleep/wake phenotyping.

High-performance in the SSS system enabled us to obtain various sleep/wake phenotypes based not only on daily sleep time (or awake time) but also on the transition probabilities between sleep and awake states ( $s$  and  $w$ , respectively, in  $P_{ws}$  and  $P_{sw}$ ; Figure 3F) and on the sleep/wake amplitude, which is related to the sleep-time variation within the daily sleep cycle (Figure 3G; Supplemental Experimental Procedures). We note that  $P_{ws}$  and  $P_{sw}$  were proposed to represent sleep drive and wake drive in a previous study (De Valck and Cluydts, 2003) and that they are informative sleep/wake phenotypes. We also defined the sleep/wake amplitude as the coefficient of variation (CV, the SD divided by the mean) of sleep time for each 10-min bin for 24 hr (Figure 3G).

To evaluate the performance of the SSS system, we recorded 108 individual C57BL/6N male mice (6 weeks old) in the SSS chamber for a week (Figure 3H; Table S1). To avoid the first-day effect, the recorded data from the second day to the seventh day were analyzed by the SSS annotation pipeline. The results revealed a daily sleep time of  $735.7 \pm 50.8$  min (mean  $\pm$  SD;  $n = 108$ ), which corresponded to a CV of 6.9%. The  $P_{ws}$  and  $P_{sw}$  were  $0.0379 \pm 0.0068$  and  $0.0363 \pm 0.0067$ , respectively,

### Figure 2. Exome Analysis of Whole-Body Biallelic *Tyr* KO Mice Produced by Triple-Target CRISPR Method

Genome alignments of exome sequence reads. The genomic regions targeted by the three gRNAs were shown for wild-type, *Tyr* KO #1, and *Tyr* KO #2. The upper part of each panel shows read coverage, and the lower part shows read alignments (rectangles). Horizontal lines between rectangles indicate read pairing. Different types of mutations occurring at the target sites are highlighted as follows: inter-exon deletion and inversion (red rectangles); intra-exon deletion and inversion mutation (green rectangles); short deletion (red dots); and short insertion (black dots).



**Figure 3. The SSS Enables Non-invasive, Fully Automated, High-Performance Sleep/Wake Phenotyping**

(A) The SSS system. Upper left: one mouse is housed in an SSS chamber. The chamber is connected to a water outlet, and food is stored in its ceiling, which allows 14 days of continuous recording without additional handling of the chamber. Upper right: up to 36 SSS chambers are set in an SSS rack. The rack is

(legend continued on next page)

and the sleep/wake amplitude was  $0.4462 \pm 0.0530$ . To evaluate the SSS system for other inbred strains and in females, we recorded C57BL/6J male mice, C57BL/6N female mice, and C57BL/6J female mice ( $n = 178, 72, \text{ and } 173$ , respectively; [Figure S2E](#); [Table S1](#)). Their tight distributions of sleep time showed that the SSS provides a simple and high-performance method for sleep/wake phenotyping.

To further validate the SSS, we recorded various known sleep/wake phenotypes in mice. First, we tested mice whose sleep/wake behavior was perturbed pharmacologically. Methylamphetamine (MAP), a potent psychostimulant that promotes wakefulness ([Edgar and Seidel, 1997](#)), was injected intraperitoneally (i.p.) into mice during the day, at zeitgeber time 2 (ZT2; [Figure S3A](#)). The administration of 3 mg/kg MAP significantly decreased sleep time and  $P_{ws}$  ([Figures S3B and S3C](#)). These results indicated that the decreased sleep time was due to increased stabilization of the awake state rather than by decreased stabilization of the sleep state.

To test the SSS system for the opposite phenotype (sleep-time increase), we injected i.p. a sleep-inducing drug, diazepam (DZP) ([Radulovacki et al., 1984](#)), at nighttime, ZT13 ([Figure S3D](#)). Indeed, the administration of 5 mg/kg DZP significantly increased the sleep time and  $P_{ws}$  ([Figures S3E and S3F](#)), which is the exact opposite phenotype observed in MAP injection. To further confirm observed sleep-phenotype changes, we performed the drug administration experiment under simultaneous recording of SSS and EEG/EMG ([Figures S3H and S3I](#)). The accuracy, sleep sensitivity, sleep specificity, wake sensitivity, and wake specificity of SSS against EEG/EMG-based sleep staging were as high as those for wild-type animals ([Figures 3D and 3E](#)). These results clearly show that SSS is highly sensitive to sleep/wake pattern, even in animals under drug administration.

Furthermore, to test the SSS system's ability to detect the sleep/wake phenotypes of genetically modified animals, we recorded the behavior of circadian mutants that lose their circadian rhythm under constant darkness (DD) ([Laposky et al., 2005](#); [Shir-omani et al., 2004](#); [Wisor et al., 2002](#)). Three KO strains, *Bmal1* KO ([Shimba et al., 2011](#)), *Cry1/Cry2* ([van der Horst et al., 1999](#)), and *Per1/Per2* double-KO mice ([Bae et al., 2001](#)) were re-

corded for a week under LD conditions (light on at ZT0, light off at ZT12; [Figure S3J](#)). All strains exhibited less sleep during the light phase (an inactive phase of wild-type mice) and more sleep during the dark phase (an active phase for wild-type mice; [Figure S3K](#)). Both  $P_{ws}$  during the light phase and  $P_{sw}$  during the dark phase are decreased in *Bmal1* KO and *Cry1/Cry2* double-KOs ([Figures S3L and S3M](#)). Although *Per1/Per2* double-KO mice did not exhibit a significant decrease of sleep/wake amplitude, their daily sleep/wake patterns differed from those of wild-type mice ([Figure S3K](#)). In all strains, the total sleep time was not significantly different from that of the C57BL/6N mice ([Figure S3M](#); [Table S1](#)). Overall, these results also indicate that the SSS can sensitively detect various sleep/wake phenotypes, including the total sleep/wake time, transition probabilities between sleep and awake states, and sleep/wake amplitude.

### Combined Use of a Triple-Target CRISPR and SSS Enables Sleep/Wake Phenotyping, Even in Double-KO Mice

To test the feasibility of combining efficient KO mouse production with automated sleep/wake phenotyping, we analyzed sleep/wake phenotypes in different KO mouse lines produced by the triple-target CRISPR method described earlier. First, we designed three gRNAs against the *Bmal1* gene ([Figure S4A](#)). Sleep/wake phenotyping was performed for animals with the KO genotype, which was confirmed by qPCR and/or genomic sequencing ([Figures S4D and S4E](#); [Tables S1 and S2](#)). As expected, the generated *Bmal1* KO mice showed a significantly decreased sleep/wake amplitude under LD conditions compared with wild-type mice, and the *Tyr* KO mice did not ([Figures 4A–4C](#); [Figure S4H](#)). In contrast, the daily sleep times did not differ between the *Tyr* KO and *Bmal1* KO mice or in comparison with wild-type mice ([Figure 4C](#)). Notably, the *Tyr* KO mice had sleep/wake patterns similar to those of the wild-type mice ([Figure 4B](#)), indicating that the triple-target CRISPR method itself had little effect on sleep/wake phenotypes. Next, we generated double-KO mice for *Cry1/Cry2* and *Per1/Per2*, which are well-known mouse clock genes. We designed three gRNAs each for each gene ([Figures S4B and S4C](#)), and sleep/wake

equipped with continuous air circulation units, which provides fresh air to each chamber. Lower panel: we used six SSS racks for recording and thus could determine the sleep/wake phenotypes of 216 mice per week.

(B) Equivalent electric circuit model of the SSS system. The SSS chamber (red dashed box) was modeled with a capacitor ( $C_1$ ) and resistor ( $R_1$ ) in parallel, and the pressure sensor (blue dashed box) was modeled with the resistor ( $R_2$ ) and capacitor ( $C_2$ ) in series ([Supplemental Experimental Procedures](#)).

(C) Calculation of the conductance over  $R_2$ , which corresponds to the amplification factor of the pressure detected with the sensor and optimization of  $C_2$  and  $R_1$  to maximize the gain of the frequency band of the physiological respiration of mice (1 to 10 Hz, pink shaded area). Based on the simulation results,  $C_2$  was fixed to  $6.12 \times 10^{-5}$  ml/Pa, and  $R_1$  was fixed to 172 Pa · s/ml.

(D) SSS-based sleep/wake staging results compared with conventional EEG/EMG-based sleep/wake staging results. Data from three mice were used. Upper graph shows the sleep time for every 30 min evaluated by SSS. Lower graph shows the accuracy based on conventional EEG/EMG-based staging, for each 30 min during recording.

(E) The performance of SSS. Sleep (red) and awake (blue) states of three individual mice are shown as probability densities on a two-dimensional principal-component map in the SSS annotation pipeline. Values represent the mean  $\pm$  SD ( $n = 3$ ) of the sensitivity and specificity of sleep (red) and awake (blue) states when compared with EEG/EMG-based sleep/wake staging. The overall accuracy reached  $95.3\% \pm 0.4\%$  (mean  $\pm$  SD, green).

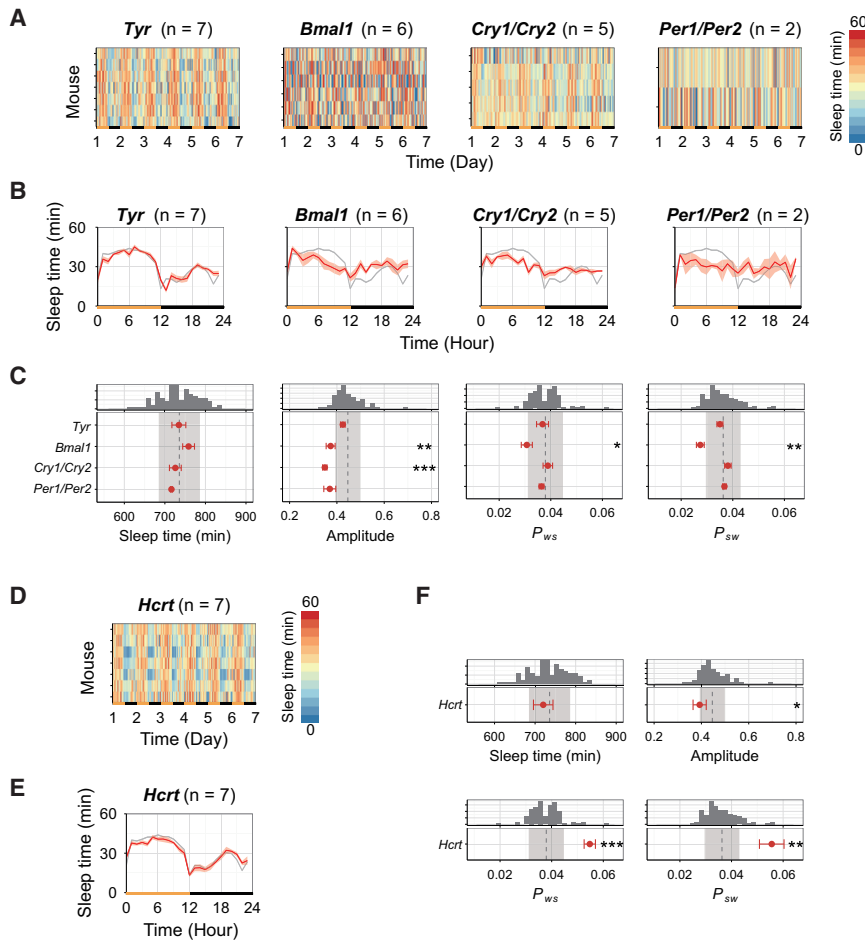
(F) Illustration of sleep/wake transitions. Upper left: schematic view of the transition probabilities between sleep and awake states. Upper right: sleep time at different values of  $P_{ws}$  and  $P_{sw}$ . When both  $P_{ws}$  and  $P_{sw}$  decrease (green circle), both sleep and awake episodes will stabilize. Note that when both  $P_{ws}$  and  $P_{sw}$  increase (red diamond), the stability of each state is dramatically decreased (bar chart below) without changing the total sleep (or awake) time.

(G) Illustration of sleep/wake amplitude. In the chart, different amplitudes are distinguished by color, and the points represent sleep times at various times of day.

(H) Distribution of sleep/wake parameters of C57BL/6N male mice recorded in the SSS. The pink and sky blue dashed lines show the mean and the mean  $\pm$  SD of each distribution, respectively.

See also [Table S1](#) and [Supplemental Experimental Procedures](#).





**Figure 4. Combined Use of a Triple-Target CRISPR and SSS Enables Sleep/Wake Phenotyping in Single- and Double-KO Mice**

(A) Sleep time (per hour) phenotype for individual mice in *Tyr* KO mice and circadian-clock mutants produced by triple-target CRISPR.

(B) Sleep time (per hour) over 24 hr averaged over 6 days in *Tyr* KO mice and circadian-clock mutants produced by triple-target CRISPR. Red lines show the mean sleep time at each time of day for each strain. Gray line shows the data for C57BL/6N male mice ( $n = 108$ ). The shaded area around the line is the SEM for each time point.

(C) Distributions of sleep/wake parameters of *Tyr* KO mice and circadian-clock mutants. Dunnett's test compared to C57BL/6N male mice.

(D) Sleep-time (per hour) phenotype of *Hcrt* KO mice.

(E) Sleep time (per hour) over 24 hr averaged over 6 days in *Hcrt* KO mice. Red line, gray line, and the shaded area are the same as described in (B).

(F) Distributions of sleep/wake parameters of *Hcrt* KO mice. Dunnett's test compared to C57BL/6N male mice. See also Table S1.

phenotyping was performed for animals with the KO genotype (Figures S4F and S4G; Tables S1 and S2). Both strains showed similar phenotypes with conventionally generated KO mice (Figures 4A–4C; Table S1). All three circadian KO strains were also recorded for a week under DD conditions and confirmed to exhibit arrhythmic phenotype (Figures S4I and S4J).

To further validate the functionality of our strategy, we designed three gRNAs against the *Hcrt* gene (Figure S4K) and produced animals by triple-target CRISPR method. Sleep/wake phenotyping was performed for animals with KO genotype (Figure S4L; Tables S1 and S2). Consistent with the conventional *Hcrt* KO mice (Chemelli et al., 1999), the mice showed significant increase in both transition probabilities,  $P_{ws}$  and  $P_{sw}$ , without any significant change in total sleep time (Figures 4D–4F and S4M).

Overall, the combinatorial use of triple-target CRISPR and SSS enables the efficient generation of KO mice and their accurate sleep/wake phenotyping not only in single-KO mice but also in double-KO mice.

### Reverse Genetics without Crossing Reveals *Nr3a* Mutant to Be a Short-Sleeper Mouse

Mammalian reverse genetics without crossing can be applied to the comprehensive analysis of gene functions at the systems and whole-organism level. To demonstrate proof of concept,

we comprehensively targeted the members of the NMDAR family. We focused on this family because the NMDAR is implicated in several psychological disorders, including schizophrenia (Belforte et al., 2010; Moghaddam and Javitt, 2012; Mohn et al., 1999), Alzheimer's disease (Snyder et al., 2005), and depression (Autry et al., 2011; Li et al., 2010), all of which have associated sleep disorders.

There are seven genes in the NMDAR family: *Nr1*, *Nr2a*, *Nr2b*, *Nr2c*, *Nr2d*, *Nr3a*, and *Nr3b*. For triple-target CRISPR, we designed three gRNAs for each of these genes (Figures S5A–S5G) and used them to generate KO mice for each one. Sleep/wake phenotyping was performed for animals with the KO genotype, which was confirmed by qPCR and/or genomic sequencing (Figures S5H–S5L; Tables S1 and S2). First, we confirmed that the lethal phenotype seen in *Nr1* and *Nr2b* biallelic KO mice that was generated by conventional methods (Forrest et al., 1994; Li et al., 1994) also occurred in *Nr1* and *Nr2b* KO mice generated by the triple-target CRISPR method, because we noted that there were no survivors of *Nr1* and *Nr2b* KO genotypes (Table S2). KO mice for the other five NMDAR genes, *Nr2a*, *Nr2c*, *Nr2d*, *Nr3a*, and *Nr3b*, survived for at least 6 to 7 weeks (Table S2), at which time they were placed in SSS chambers, and their sleep/wake behavior was recorded for at least 1 week under LD conditions. For the phenotype analysis, we used the 6-day data from the second recording day to avoid the first-day effect (Figure 5A). The sleep/wake phenotype analysis was performed fully automatically, and the statistical evaluation used sleep time,  $P_{ws}$ ,  $P_{sw}$ , and sleep/wake amplitude as parameters (Figures 5B, 5C, and S5M; Table S1).

Among the five KO mice, the *Nr3a* KO mice showed a significant “short-sleeper” phenotype, with a daily sleep time of

639.0 ± 17.5 min (n = 8), which is 96.7 min shorter than that of wild-type mice (p < 0.0001). *Nr3a* KO mice displayed a lower  $P_{ws}$  than wild-type mice, whereas their  $P_{sw}$  was normal (Figure 5C), indicating that their short-sleeper phenotype was due to an increase in awake-state stabilization (i.e., longer episode durations of the awake state) but not a decrease in sleep-state stabilization. To preclude the possibility of an off-target effect of CRISPR/Cas9 gene editing, we generated another group of *Nr3a* KO mice (n = 12) by an independent CRISPR probe set (set 2; Figures S6A and S6B; Table S2) and confirmed the observed short-sleeper phenotype compared with wild-type mice (p < 0.0001) (Figures 5D–5F and S6C). Furthermore, to ensure that the short-sleeper phenotype is not due to the modality of sleep evaluation, we produced another series of *Nr3a* (set 1) KO mice (Figure S6D) and used EEG and EMG recordings to evaluate their sleep. Consistent with the SSS results, EEG recordings showed a significant short-sleeper phenotype in KO mice (p < 0.05; Figure S6E). For genomic validation, four of the generated *Nr3a* KO mice (Figures S5K and S6B) mentioned earlier were analyzed with exome sequencing (Table S5). In both set 1 and set 2 animals, the analysis revealed deletions of genome stretches with variable lengths at the target site (Figure 6; Figure S7A), with its extent for at least one target site of up to 100% (see Figure S7B for a magnifiable view). Importantly, no off-target mutation was identified in this exome sequencing. These results strongly suggested that the observed short-sleeper phenotype of *Nr3a* KO mice cannot be attributed to the possible off-target effect of CRISPR but rather to the common genomic defects in the *Nr3a* gene. Finally, to confirm the absence of NR3A protein in the KO animals, we have performed MS analysis in these animals (Figure S6F). Two *Tyr* KO animals (Figure S4D) were used as controls for NR3A protein detection, which showed a clear peak at the position of internal standard of NR3A protein. By contrast, none of the *Nr3a* KO mice (Figures S5K and S6B) had a detectable peak at the position of internal NR3A protein. These results suggest that *Nr3a* is knocked out not only at the genomic level but also at the protein expression level. Overall, these results demonstrate that the use of mammalian reverse genetics without crossing, and specifically combining this efficient KO mouse production with accurate, automated phenotyping, enables the comprehensive identification of shared and diverse functions of gene family members in organisms.

### Developing a Database for the Triple-Target CRISPR Method

As described earlier, triple-target CRISPR enabled us to efficiently produce biallelic KO mice. However, as we produced several KO mice using this method, we noticed that manually designing triple targets for the same gene is still time consuming, even using existing tools to extract candidate target sequences for a given genomic sequence (<http://cas9.cbi.pku.edu.cn/index.jsp>) (Ma et al., 2013) and to evaluate the off-target risk for each candidate (<http://tools.genome-engineering.org>) (Ran et al., 2013). This manual design procedure takes up to a few hours for a single gene and, therefore, limits the ability to perform large-scale screenings. Therefore, we sought to develop an automated method that performs all of the gRNA selection steps.

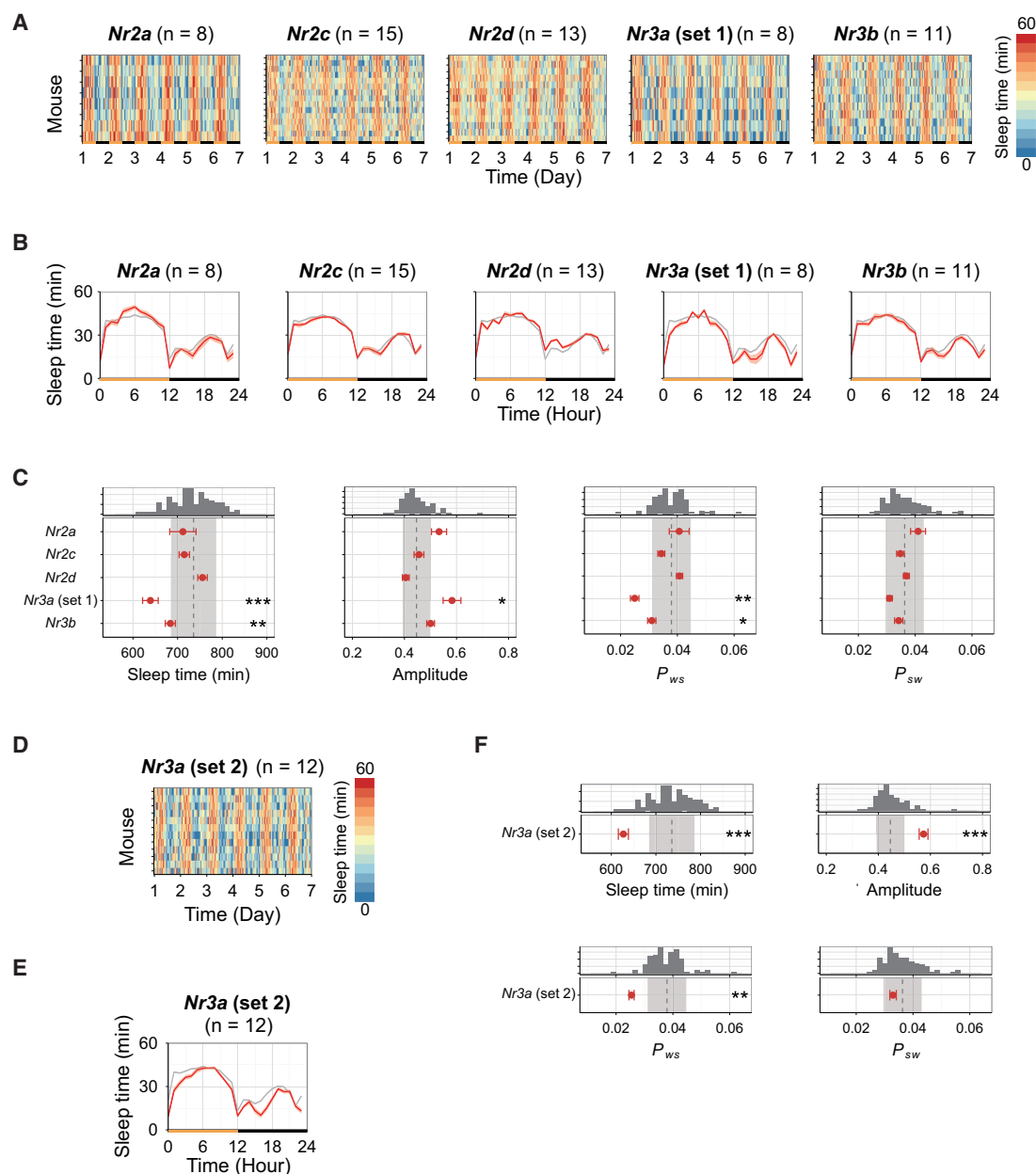
We scanned the whole mouse genome once, and all of the suitable targets are now in an online database (Figure 7A; Supplemental Experimental Procedures). This gRNA database provides at least three target sequences (one set for triple-target CRISPR) for 81.2% of mouse genes. In addition, 71.9% of all mouse genes have more than six target sequences (multiple sets for triple-target CRISPR), which are also included (Figures 7B and 7C).

To validate the usefulness of the triple-target CRISPR database, we selected two independent sets of triple targets for the *Tyr* gene (Tyr-4, Tyr-5, and Tyr-6 for set 2, and Tyr-7, Tyr-8, and Tyr-9 for set 3; Figure 7D), which were distinct from Tyr-1, Tyr-2, and Tyr-3 (set 1) used in the previous experiments (Figure 1C). When injected into fertilized eggs, the triple-target method using both sets of gRNAs showed almost perfect efficacy (96.6% and 100%, respectively; Figures 7E and 7F), supporting the usefulness of the triple-target CRISPR database. Therefore, we made this database public (<http://crispr.riken.jp/>).

## DISCUSSION

### Triple-Target CRISPR Enables the Efficient Production of Biallelic KO Mice in a Single Generation

In this study, we sought to achieve the highly efficient (>90%) production of whole-body biallelic KO mice in a single generation. We used a simple computational model to estimate the minimum efficacy of different CRISPR methods and discovered that multiple-target CRISPR methods would generate mutant mice with much higher efficiency than single-target CRISPR. We designed three different gRNAs against the *Tyr* gene and confirmed that the triple-target CRISPR method—using a mixture of three gRNAs (Tyr-1, Tyr-2, and Tyr-3)—achieved almost perfect efficacy (97.5%), whereas the single-target method with any of the gRNAs at a 3-fold higher concentration had only moderate efficiency. Genomic verification revealed no off-target mutation, at least in exons (Figure 2; Table S5). The total number of gigabase pairs (or Gbp) read in the exome sequencing including *Tyr* (set 1), *Nr3a* (set 1), and *Nr3a* (set 2) KOs was 65.6 Gbp. This is more than 24 times greater than the whole genome base pairs of mouse (2.7 Gbp). This, therefore, gives us enough coverage to be confident that off-target effect of the triple-target CRISPR strategy is not a problem in practice. Importantly, the DNA-cleavage efficiency of Tyr-3 was more than 3-fold higher than that of Tyr-2 (Figure 1D), whereas the whole-body biallelic KO efficiency for Tyr-3 (64.7%) was similar to that for Tyr-2 (54.2%) (Figure 1F), indicating that a limiting factor in vivo is not DNA cleavage by the CRISPR/Cas9 system but other factors such as DNA repair, consistent with the prediction by the computational model. The reproducibility and robustness of the triple-target method were further confirmed by additional experiments with two independent sets of three gRNAs against the *Tyr* gene (Figures 7E and 7F). We also confirmed the KO efficiency by genotyping all mice used for SSS analysis in this study. We genotyped 102 mice produced by the triple-target CRISPR method, including 11 single-KO and 2 double-KO strains (Table S2). At least 92.2% (n = 94) of the animals were confirmed as KO mice by qPCR or sequencing. This efficacy is confirmed to be higher than the minimum efficacy predicted by our model. In addition, the triple-target CRISPR method is scalable and can be applied to the identification of common and

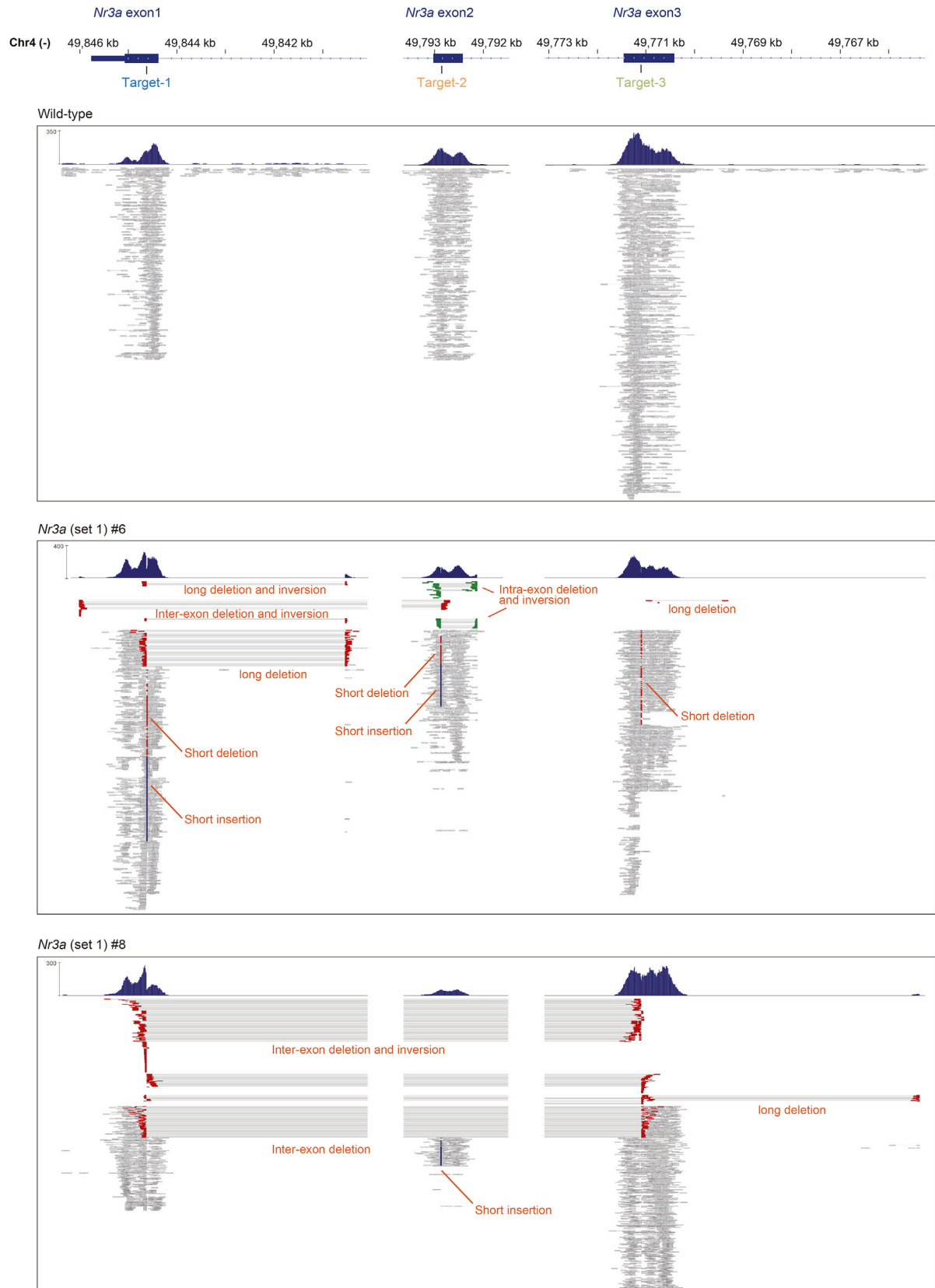


**Figure 5. Reverse Genetics without Crossing Reveals *Nr3a* Mutant to Be a Short-Sleeper**

(A) Sleep-time (per hour) phenotype for all individual KO mice of five NMDA receptor family members generated by triple-target CRISPR. (B) Sleep time (per hour) over 24 hr, averaged over six days in KO mice of five NMDA receptor family members. Red lines show the mean sleep time at each time of day for each strain. Gray line shows the data for C57BL/6N male mice ( $n = 108$ ). The shaded area around the line is the SEM for each time point. (C) Distributions of sleep/wake parameters of comprehensive KO mice of NMDA receptor family members. Dunnett's test or Dunnett's modified Tukey-Kramer pairwise multiple comparison test compared to C57BL/6N male mice. (D) Sleep-time (per hour) phenotype of *Nr3a* KO mice (set 2) generated by triple-target CRISPR. (E) Sleep time (per hour) over 24 hr, averaged over 6 days in *Nr3a* KO mice (set 2). Red line, gray line, and the shaded area are the same as described in (B). (F) Distributions of sleep/wake parameters of *Nr3a* KO mice (set 2). Dunnett's test or Dunnett's modified Tukey-Kramer pairwise multiple comparison test compared to C57BL/6N male mice. See also [Table S1](#).

diverse functions of a gene family of interest (Figure 5). Furthermore, the triple-target CRISPR method is applicable to the phenotype analysis of double-KO mice (Figures 4A–4C). We also

created a publicly available CRISPR database that provides triple-target candidate gRNAs for more than 81.2% of the genes in the mouse genome (Figure 7A).



(legend on next page)

### The SSS Provides a Simple and Accurate Phenotyping Method for a Complex Behavior

Along with the efficient production of biallelic KO mice, accurate phenotype analysis is also important to make full use of mammalian reverse genetics without crossing. In this study, we focused on the phenotype analysis of sleep/wake states, because sleep disorders are often associated with psychiatric diseases. The development of a non-invasive and fully automated method for sleep/wake phenotyping (the SSS) allowed us to sensitively detect the sleep-time effects of drugs targeting the CNS, including a psychostimulant, MAP (dopamine transporter antagonist), and a known sleep-inducing drug, DZP (GABA<sub>A</sub> receptor agonist) (Figures S3A–S3F). In addition, the SSS enabled us to accurately and quantitatively analyze other sleep/wake phenotypes, such as sleep/wake amplitude, which was shown by the decreased amplitude of clock mutants (Figure 4C), as well as the transition probabilities from sleep to awake states, and from awake to sleep states (Figures 3F–3H), which was clearly shown by the significant increase of  $P_{ws}$  and  $P_{sw}$  in *Hcr1* KO mice (Figure 4F). This ability is useful for dissecting the effects of CNS drugs on daily sleep time. For example, we found that the decreased sleep time caused by MAP was due to an increase in awake-state stabilization (low  $P_{ws}$ ) rather than to a decrease in sleep-state stabilization (Figure S3B). In contrast, the increased sleep time caused by DZP for 5 hr after the injection was due to a decrease in awake-state stabilization (high  $P_{ws}$ ) (Figure S3E), suggesting that the sleep/wake phenotype induced by DZP is functionally the opposite of that induced by MAP. Interestingly, after DZP increased the sleep time (0 to 5 hr after injection), it conversely decreased the sleep time (5 to 10 hr after injection) (Figure S3G). This compensatory decrease in sleep time by DZP was associated with a slight decrease in sleep-state stabilization (high  $P_{sw}$ ) (Figure S3G;  $p = 0.0869$ ). These results suggest that the SSS can sensitively detect and carefully dissect the common and diverse sleep/wake effects of drugs targeting the CNS.

The sleep time of two circadian KO animals, *Bmal1* KO and *Cry1/Cry2* double-KO mice, recorded in the SSS is not exactly the same as in past reports (Laposky et al., 2005; Wisor et al., 2002). This might be partly because of the functional advantages of the SSS system. First, because of its convenience, the number of recorded mutant mice is larger in this study, where 34 *Bmal1* KO and 27 *Cry1/Cry2* KO mice and more than 100 wild-type mice are recorded. In contrast, nine *Bmal1* KO mice against eight wild-type mice are recorded in the Laposky et al. (2005) paper, and six *Cry1/2* double-KO mice against eight wild-type mice are recorded in the Wisor et al. (2002) paper. Second, in both studies, the sleep phenotypes were evaluated by electrodes implanted in the skull, which is inevitably invasive. In fact, three out of nine KO mice were removed from the study after EEG/EMG recording in the Laposky et al. paper because EEG/EMG signals were deteriorating,

which implies some invasiveness inherited in the conventional EEG/EMG recording methods.

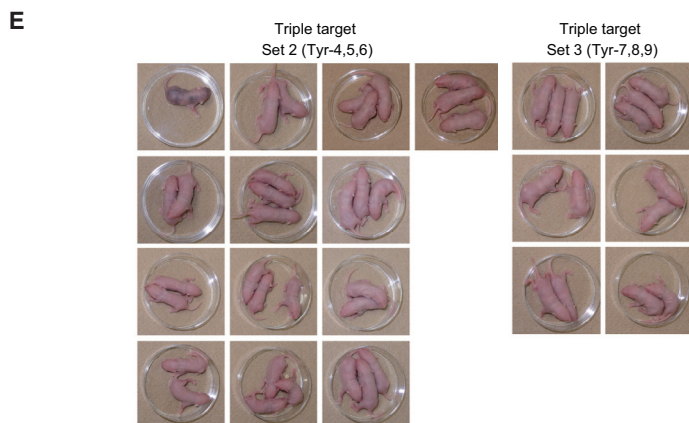
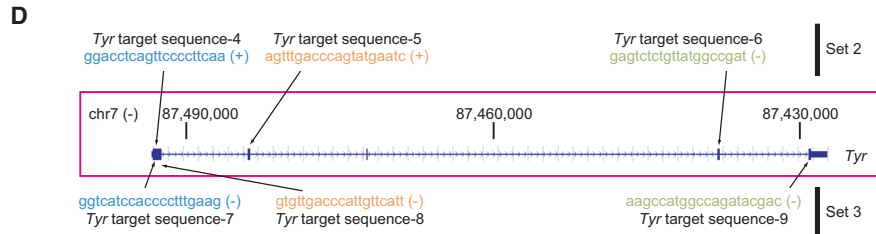
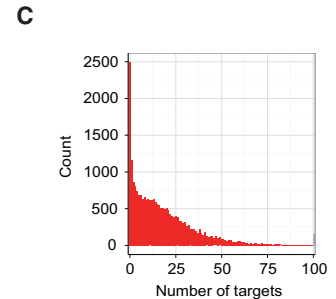
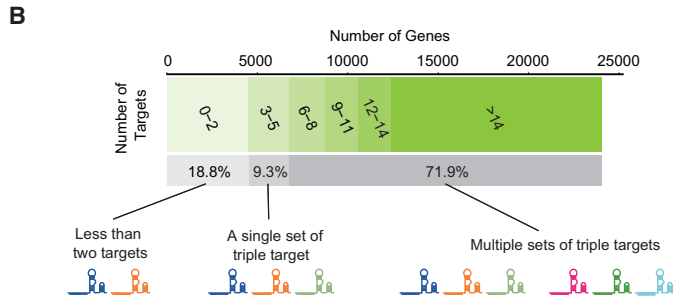
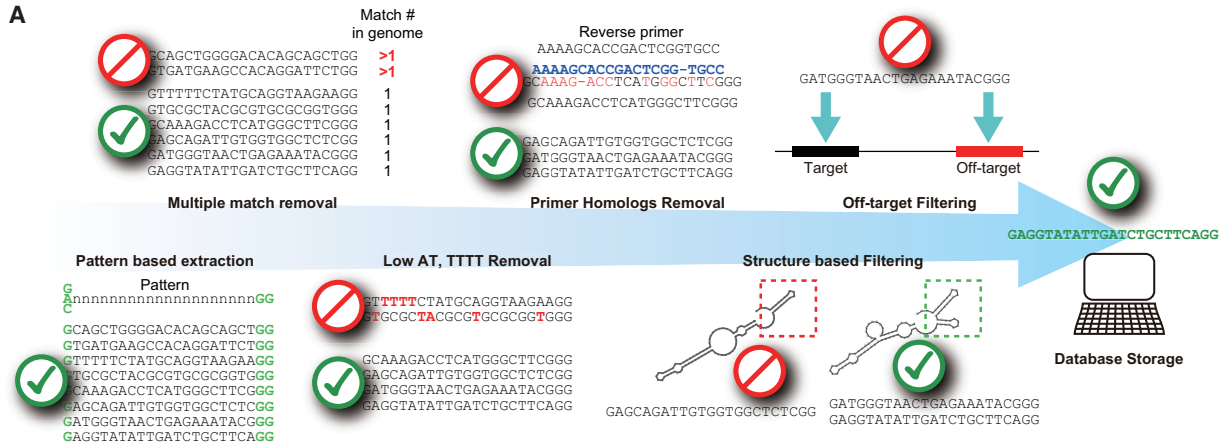
Several other approaches have been used in the past to develop automated non-invasive sleep staging systems. The video-based sleep stager is non-invasive and can discriminate between REM and NREM (McShane et al., 2012). However, the expected drawback is difficulty in parallelization due to the huge amount of data to be handled. Furthermore, this method requires a user to set a parameter, which makes it distinct from the SSS, which requires no manual optimization. Piezoelectric-based movement detection technology allowed for estimation of the sleep/wake status from the animal motion (Flores et al., 2007). The simplicity of this system is attractive, although when performing a high-throughput phenotyping in this system, we expect its inevitable interference, i.e., the difficulty in assigning the movement information to the recoding animal against the inevitable noise coming from the neighboring chambers. In contrast, the SSS only uses single-channel respiration flow information per animal, which is a very compact-sized dataset and the respiration is detected only by the pressure inside the chamber, which is much less prone to be affected from the neighboring chambers. These characteristics make the SSS a much useful technology.

### Regulation of Sleep/Wake States by NMDAR Family Members

To test the use of our combined method to comprehensively identify the common and diverse functions of gene family members, we separately knocked out all of the NMDAR family members. We found that *Nr3a* KO mice are short-sleeper mutants. *Nr3a* expression peaks during early postnatal life and decreases to lower levels in adulthood (Henson et al., 2010). Unlike the traditional NMDARs, *Nr3a*-containing NMDARs have less Ca<sup>2+</sup> permeability and are almost completely insensitive to Mg<sup>2+</sup> block at hyperpolarized potentials (Sasaki et al., 2002; Tong et al., 2008). Recently, adult *Nr3a* KO mice were reported to have increased pain sensation and enhanced abilities in learning and memory tasks (Mohamad et al., 2013). These KO mice had greater long-term potentiation (LTP) in adult hippocampal slices compared with wild-type slices. Based on these previous findings, we can assume two possible hypotheses to explain the short-sleeper phenotype in *Nr3a* KO animals. One is the developmental circuit influence of *Nr3a*-containing neurons. Since *Nr3a* defect impairs fundamental NMDAR responses that are deeply related to formation of synaptic networks, the absence of *Nr3a* may alter the original sleep regulation network of the brain. The other is the direct neural activity modulation by NR3A-containing neurons. Even *Nr3a* impairment during development may be complemented by another mechanism; the influence of *Nr3a* modulation against neural activity remains until adulthood. The enhanced LTP can be explained by decreased blunting of the NMDAR response by *Nr3a*. Because the NMDAR

#### Figure 6. Exome Analysis of Whole-Body Biallelic *Nr3a*, Set 1, KO Mice Produced by Triple-Target CRISPR Method

The genomic regions targeted by the three gRNAs were shown for two mice in *Nr3a* (set 1) KO group (Figure S5K). The upper part of each panel shows read coverage, and the lower part shows read alignments (rectangles). Horizontal lines between rectangles indicate read pairing. Different types of mutations occurring at the target sites are highlighted as follows: inter-exon deletion (red rectangles); intra-exon deletion mutation (green rectangles); short deletion (red dots); and short insertion (blue dots).



**F**

	Triple Set 2 (Tyr-4,5,6)	Triple Set 3 (Tyr-7,8,9)
gRNA1 (ng/μl)	50	50
gRNA2 (ng/μl)	50	50
gRNA3 (ng/μl)	50	50
Embryos	145	105
Pups	32	14
Albino	31 (96.9%)	14 (100%)
Pigmented	1 (3.1%)	0 (0%)

(legend on next page)

family can be strongly blocked by MK-801, which results in hyperlocomotion when administered to mice (Koek et al., 1988), it is possible that, without *Nr3a*, the NMDAR response is weakened, and this may result in shorter sleep. To test these hypotheses, conditional perturbation of *Nr3a*-containing neurons is necessary in future investigations.

Although the detailed molecular mechanism, in which *Nr3a* contributes to the sleep/wake cycle, remains to be elucidated, the discovery of *Nr3a* KO mice as a strong short sleeper provides an interesting entry point for mammalian sleep research. We also note that, in the fruit fly, forward genetic screenings have shown that other ion channels play important roles in sleep homeostasis. Mutation in the *Shaker* gene, which encodes a voltage-dependent potassium channel, and in the *Redeye* gene, which encodes a nicotinic acetylcholine receptor  $\alpha$  subunit, both result in a short-sleeping phenotype (Cirelli et al., 2005; Shi et al., 2014). Therefore, it will be interesting to investigate the relationship of the mammalian homologs of these channels to NR3A-containing NMDARs in mammalian sleep.

### Toward Organism-Level Systems Biology

In this study, we sought to develop mammalian reverse genetics without crossing by eliciting the efficient (>90%) production and accurate (>90%) phenotype analysis of mutant animals in a single generation. Mammalian reverse genetics without crossing (called “next-generation” mammalian reverse genetics) now allows us to test—in whole organisms—hypotheses that are derived from the past knowledge. For example, a computational model of slow-wave sleep that provides explanations for the observed complex dynamics of sleep/wake states, along with predictions for the important molecular pathways regulating these states, could be extremely powerful. Previously, it was difficult to test such predictions, at least without years of labor. However, the present study demonstrated that the importance of a predicted molecular pathway can be efficiently tested using next-generation reverse genetics. Therefore, one of the next challenges will be to construct a simple computational model and experimentally validate its predictions using triple-target CRISPR and/or the SSS. New research fields, including the system-level identification and analysis of molecular circuits in whole organisms, lie ahead.

### EXPERIMENTAL PROCEDURES

#### Animals

C57BL/6NJcl mice were purchased from CLEA Japan, and C57BL/6J mice were purchased from Oriental Yeast. All mice were given food and water ad libitum. They were kept in an environment at an ambient temperature of 21°C, with a relative humidity of 50%. The light was controlled under a

12-hr/12-hr light/dark cycle. All experimental procedures and housing conditions involving animals and their care were in accordance with the RIKEN Regulations for Animal Experiments and the Animal Care and the Use Committee of the Graduate School of Medicine, The University of Tokyo.

#### One-Cell Embryo Microinjection

C57BL/6N females (4–6 weeks old) were superovulated and mated with C57BL/6N males. Fertilized eggs were collected from the ampulla of the oviduct of plugged C57BL/6N females by micro-dissection and kept in KSOM medium (Merck Millipore) in a 5% CO<sub>2</sub> incubator at 37°C. Cas9 mRNA (100 ng/ $\mu$ l) and gRNAs (150 ng/ $\mu$ l in total) were co-injected into the cytoplasm of fertilized eggs in M2 medium (Merck Millipore) at room temperature. Details of the cytoplasmic injection were described previously (Sumiyama et al., 2010). After microinjection, the injected embryos were cultured for 1 hr in KSOM medium (Merck Millipore) in a 5% CO<sub>2</sub> incubator at 37°C, and 15–30 embryos were then transferred to the oviducts of pseudopregnant ICR female mice.

#### Non-invasive Sleep/Wake Phenotype Analysis by SSS Recordings

Mice were placed in the SSS chamber (Figure 3A, upper left panel; Minamide Rika Shokai), which had an air flow of at least 0.4 l/min. The chamber was connected to the SSS sensor (PMD-8203-2G, Biotex, light brown box on top of the chamber in Figure 3A), which detected the pressure difference between the outside and inside of the chamber. The analog signal output from the sensor was digitized at 250 Hz by an AD converter (ADS-0128AK-64MB, Microscience) and stored on a computer by data logging software (Lab-DAQ Pro, Matsuyama Advance). The analysis of the digitized data was fully automated as follows: first, the respiration time domain data were converted to frequency domain data by fast Fourier transformation. The power spectrum was then subjected to principal-component analysis. The first three principal components were used for clustering by probability density clustering (Azzalini and Torelli, 2007). Finally, the clusters were annotated according to the power spectrum of each cluster. The details in each step of SSS annotation pipeline are described in the Supplemental Experimental Procedures. One set of data recording in the SSS chamber was for 2 weeks. During the experiment, neither the chamber nor the rack was opened unless the injection was performed i.p. The light condition of the SSS rack was set to a 12-hr/12-hr light/dark cycle, and food and water were given ad libitum. Data analysis was conducted 6 days from the second recording date.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, five tables, and a data file showing the source code of the SSS analysis and can be found with this article online at <http://dx.doi.org/10.1016/j.celrep.2015.12.052>.

### AUTHOR CONTRIBUTIONS

H.R.U. and G.A.S. designed the study. G.A.S. developed the SSS system and performed the sleep data analysis. G.A.S. and S.S. performed the SSS recording. G.A.S., D.T., and K.L.O. performed EEG/EMG recordings. K.S. developed the triple-target CRISPR method. M.U.-T. designed the CRISPR probes. M.U.-T. and S.S. produced CRISPR probes. K.S. and R.O. produced the mice. D.P. implemented the whole-genome detection method and the

### Figure 7. A Publicly Available Database for the Triple-Target CRISPR Method

- (A) The pipeline for automatically selecting targets for the triple-target CRISPR method.  
 (B) Distribution of the number of the targets per gene.  
 (C) Histogram of the number of targets per gene.  
 (D) Alternative target sequences for the *Tyr* gene.  
 (E) Experimental validation of the automatically selected targets for the *Tyr* gene. Two sets of additional triple gRNAs (Tyr-4, Tyr-5, and Tyr-6 for set 2 and Tyr-7, Tyr-8, and Tyr-9 for set 3) were tested. Photographs show the coat color of the generated mice.  
 (F) Table shows the injection conditions of the gRNAs and resulting phenotype. Embryos, number of injected and transferred embryos; Pups, number of born pups (total); Albino, pups with albino coat color (biallelic knockout); Pigmented, pups with mosaic or wild-type coat color.

database for the triple-target CRISPR method. H.U. developed the SSA assay. M.U.-T. and H.F. performed the genotyping of animals. O.N. and S.K. performed the exome sequencing analysis. R.N. and Y.S. performed the protein quantification. H.R.U. and G.A.S. wrote the manuscript. All authors discussed the results and commented on the manuscript text.

## ACKNOWLEDGMENTS

We thank T. Mikami for the enthusiastic support for developing the SSS. We also thank all the lab members at RIKEN CDB, QBIC, and The University of Tokyo, in particular, S. Fujino, J. Hara, J. Yoshida-Garcon, M. Shiokawa, A. Kishimoto, A. Nishiyama, A. Sato, Y. Wada, and M. Nomura for their kind help in preparing the materials and supporting experiments; the LARGE, RIKEN CDB for housing the mice; Kaori Tatsumi in the Phyloinformatics Unit (GRAS), RIKEN Center for Life Science Technologies, for exome library preparation and sequencing; T. Ueno for running the initial experiment for triple-target CRISPR; F. Tatsuki and G. Kanda for helpful discussion; and K. Yoshida for improving the SSS algorithm. This work was supported by the Innovative Cell Biology by Innovative Technology Program from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan (H.R.U.); a Grant-in-Aid for Scientific Research (S) (grant no. 25221004; H.R.U.) and a Grant-in-Aid for Scientific Research on Innovative Areas (grant no. 23115006) from MEXT/Japan Society for the Promotion of Science (JSPS; H.R.U.); a Grant-in-Aid for Young Scientists (B) (grant no. 26870858; G.A.S.); a Grant-in-Aid for Scientific Research on Innovative Areas (grant no. 26113720; K.S.); the strategic programs for R&D (President's discretionary fund) of RIKEN (H.R.U.); an intramural Grant-in-Aid from the RIKEN Center for Developmental Biology and RIKEN Quantitative Biology Center (H.R.U.); a grant from AMED Core Research for Evolutionary Medical Science and Technology (AMED-CREST), Japan AMED (H.R.U.), and Core Research for Evolutionary Science and Technology (CREST), Japan Science and Technology Agency (JST; H.R.U.); the Program for the Brain Mapping by Integrated Neurotechnologies for Disease Studies (Brain/MINDS) from the MEXT (H.R.U.); the Basic Science and Platform Technology Program for Innovative Biological Medicine from the Japan Agency for Medical Research and Development (H.R.U.); and by the RIKEN Foreign Postdoctoral Researcher Program (D.P.).

Received: September 27, 2015

Revised: October 29, 2015

Accepted: December 8, 2015

Published: January 7, 2016

## REFERENCES

- Autry, A.E., Adachi, M., Nosyreva, E., Na, E.S., Los, M.F., Cheng, P.F., Kavalali, E.T., and Monteggia, L.M. (2011). NMDA receptor blockade at rest triggers rapid behavioural antidepressant responses. *Nature* **475**, 91–95.
- Azzalini, A., and Torelli, N. (2007). Clustering via nonparametric density estimation. *Stat. Comput.* **17**, 71–80.
- Bae, K., Jin, X., Maywood, E.S., Hastings, M.H., Reppert, S.M., and Weaver, D.R. (2001). Differential functions of mPer1, mPer2, and mPer3 in the SCN circadian clock. *Neuron* **30**, 525–536.
- Belforte, J.E., Zsiros, V., Sklar, E.R., Jiang, Z., Yu, G., Li, Y., Quinlan, E.M., and Nakazawa, K. (2010). Postnatal NMDA receptor ablation in corticolimbic interneurons confers schizophrenia-like phenotypes. *Nat. Neurosci.* **13**, 76–83.
- Carbery, I.D., Ji, D., Harrington, A., Brown, V., Weinstein, E.J., Liaw, L., and Cui, X. (2010). Targeted genome modification in mice using zinc-finger nucleases. *Genetics* **186**, 451–459.
- Chemelli, R.M., Willie, J.T., Sinton, C.M., Elmquist, J.K., Scammell, T., Lee, C., Richardson, J.A., Williams, S.C., Xiong, Y., Kisanuki, Y., et al. (1999). Narcolepsy in orexin knockout mice: molecular genetics of sleep regulation. *Cell* **98**, 437–451.
- Cirelli, C., Bushey, D., Hill, S., Huber, R., Kreber, R., Ganetzky, B., and Tononi, G. (2005). Reduced sleep in *Drosophila* Shaker mutants. *Nature* **434**, 1087–1092.
- De Valck, E., and Cluydts, R. (2003). Sleepiness as a state-trait phenomenon, comprising both a sleep drive and a wake drive. *Med. Hypotheses* **60**, 509–512.
- Edgar, D.M., and Seidel, W.F. (1997). Modafinil induces wakefulness without intensifying motor activity or subsequent rebound hypersomnolence in the rat. *J. Pharmacol. Exp. Ther.* **283**, 757–769.
- Flores, A.E., Flores, J.E., Deshpande, H., Picazo, J.A., Xie, X.S., Franken, P., Heller, H.C., Grahn, D.A., and O'Hara, B.F. (2007). Pattern recognition of sleep in rodents using piezoelectric signals generated by gross body movements. *IEEE Trans. Biomed. Eng.* **54**, 225–233.
- Forrest, D., Yuzaki, M., Soares, H.D., Ng, L., Luk, D.C., Sheng, M., Stewart, C.L., Morgan, J.I., Connor, J.A., and Curran, T. (1994). Targeted disruption of NMDA receptor 1 gene abolishes NMDA response and results in neonatal death. *Neuron* **13**, 325–338.
- Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M., and Joung, J.K. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* **32**, 279–284.
- Fujii, W., Kawasaki, K., Sugiura, K., and Naito, K. (2013). Efficient generation of large-scale genome-modified mice using gRNA and CAS9 endonuclease. *Nucleic Acids Res.* **41**, e187.
- Glaab, T., Taube, C., Braun, A., and Mitzner, W. (2007). Invasive and noninvasive methods for studying pulmonary function in mice. *Respir. Res.* **8**, 63.
- Gordon, C.J., and Ali, J.S. (1984). Measurement of ventilatory frequency in unrestrained rodents using microwave radiation. *Respir. Physiol.* **56**, 73–79.
- Henson, M.A., Roberts, A.C., Pérez-Otaño, I., and Philpot, B.D. (2010). Influence of the NR3A subunit on NMDA receptor functions. *Prog. Neurobiol.* **91**, 23–37.
- Hernandez, A.B., Kirkness, J.P., Smith, P.L., Schneider, H., Polotsky, M., Richardson, R.A., Hernandez, W.C., and Schwartz, A.R. (2012). Novel whole body plethysmography system for the continuous characterization of sleep and breathing in a mouse. *J. Appl. Physiol.* (1985) **112**, 671–680.
- Horvath, P., and Barrangou, R. (2010). CRISPR/Cas, the immune system of bacteria and archaea. *Science* **327**, 167–170.
- Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262–1278.
- Kitano, H. (2002a). Computational systems biology. *Nature* **420**, 206–210.
- Kitano, H. (2002b). Systems biology: a brief overview. *Science* **295**, 1662–1664.
- Koek, W., Woods, J.H., and Winger, G.D. (1988). MK-801, a proposed noncompetitive antagonist of excitatory amino acid neurotransmission, produces phencyclidine-like behavioral effects in pigeons, rats and rhesus monkeys. *J. Pharmacol. Exp. Ther.* **245**, 969–974.
- Laposky, A., Easton, A., Dugovic, C., Walisser, J., Bradfield, C., and Turek, F. (2005). Deletion of the mammalian circadian clock gene *BMAL1/Mop3* alters baseline sleep architecture and the response to sleep deprivation. *Sleep* **28**, 395–409.
- Li, Y., Erzurumlu, R.S., Chen, C., Jhaveri, S., and Tonegawa, S. (1994). Whisker-related neuronal patterns fail to develop in the trigeminal brainstem nuclei of NMDAR1 knockout mice. *Cell* **76**, 427–437.
- Li, N., Lee, B., Liu, R.-J., Banasr, M., Dwyer, J.M., Iwata, M., Li, X.-Y., Aghajanian, G., and Duman, R.S. (2010). mTOR-dependent synapse formation underlies the rapid antidepressant effects of NMDA antagonists. *Science* **329**, 959–964.
- Ma, M., Ye, A.Y., Zheng, W., and Kong, L. (2013). A guide RNA sequence design platform for the CRISPR/Cas9 system for model organism genomes. *BioMed Res. Int.* **2013**, 270805.
- McShane, B.B., Galante, R.J., Biber, M., Jensen, S.T., Wyner, A.J., and Pack, A.I. (2012). Assessing REM sleep in mice using video data. *Sleep* **35**, 433–442.
- Moghaddam, B., and Javitt, D. (2012). From revolution to evolution: the glutamate hypothesis of schizophrenia and its implication for treatment. *Neuropsychopharmacology* **37**, 4–15.



- Mohamad, O., Song, M., Wei, L., and Yu, S.P. (2013). Regulatory roles of the NMDA receptor GluN3A subunit in locomotion, pain perception and cognitive functions in adult mice. *J. Physiol.* 591, 149–168.
- Mohn, A.R., Gainetdinov, R.R., Caron, M.G., and Koller, B.H. (1999). Mice with reduced NMDA receptor expression display behaviors related to schizophrenia. *Cell* 98, 427–436.
- Nagy, A. (2003). *Manipulating the Mouse Embryo: A Laboratory Manual*, Third Edition (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).
- Radulovacki, M., Sreckovic, G., Zak, R., and Zahrebelski, G. (1984). Diazepam and midazolam increase light slow-wave sleep (SWS1) and decrease wakefulness in rats. *Brain Res.* 303, 194–196.
- Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* 8, 2281–2308.
- Sasaki, Y.F., Rothe, T., Premkumar, L.S., Das, S., Cui, J., Talantova, M.V., Wong, H.K., Gong, X., Chan, S.F., Zhang, D., et al. (2002). Characterization and comparison of the NR3A subunit of the NMDA receptor in recombinant systems and primary cortical neurons. *J. Neurophysiol.* 87, 2052–2063.
- Shi, M., Yue, Z., Kuryatov, A., Lindstrom, J.M., and Sehgal, A. (2014). Identification of Redeye, a new sleep-regulating protein whose expression is modulated by sleep amount. *eLife* 3, e01473.
- Shimba, S., Ogawa, T., Hitosugi, S., Ichihashi, Y., Nakadaira, Y., Kobayashi, M., Tezuka, M., Kosuge, Y., Ishige, K., Ito, Y., et al. (2011). Deficient of a clock gene, brain and muscle Arnt-like protein-1 (BMAL1), induces dyslipidemia and ectopic fat formation. *PLoS ONE* 6, e25231.
- Shiromani, P.J., Xu, M., Winston, E.M., Shiromani, S.N., Gerashchenko, D., and Weaver, D.R. (2004). Sleep rhythmicity and homeostasis in mice with targeted disruption of mPeriod genes. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 287, R47–R57.
- Snyder, E.M., Nong, Y., Almeida, C.G., Paul, S., Moran, T., Choi, E.Y., Nairn, A.C., Salter, M.W., Lombroso, P.J., Gouras, G.K., and Greengard, P. (2005). Regulation of NMDA receptor trafficking by amyloid-beta. *Nat. Neurosci.* 8, 1051–1058.
- Sumiyama, K., Kawakami, K., and Yagita, K. (2010). A simple and highly efficient transgenesis method in mice with the Tol2 transposon system and cytoplasmic microinjection. *Genomics* 95, 306–311.
- Sunagawa, G.A., Sei, H., Shimba, S., Urade, Y., and Ueda, H.R. (2013). FASTER: an unsupervised fully automated sleep staging method for mice. *Genes Cells* 18, 502–518.
- Sung, Y.H., Baek, I.J., Kim, D.H., Jeon, J., Lee, J., Lee, K., Jeong, D., Kim, J.S., and Lee, H.W. (2013). Knockout mice created by TALEN-mediated gene targeting. *Nat. Biotechnol.* 31, 23–24.
- Tong, G., Takahashi, H., Tu, S., Shin, Y., Talantova, M., Zago, W., Xia, P., Nie, Z., Goetz, T., Zhang, D., et al. (2008). Modulation of NMDA receptor properties and synaptic transmission by the NR3A subunit in mouse hippocampal and cerebrocortical neurons. *J. Neurophysiol.* 99, 122–132.
- van der Horst, G.T., Muijtjens, M., Kobayashi, K., Takano, R., Kanno, S., Takao, M., de Wit, J., Verkerk, A., Eker, A.P., van Leenen, D., et al. (1999). Mammalian Cry1 and Cry2 are essential for maintenance of circadian rhythms. *Nature* 398, 627–630.
- Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., and Jaenisch, R. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* 153, 910–918.
- Wiedenheft, B., Sternberg, S.H., and Doudna, J.A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 482, 331–338.
- Wisor, J.P., O'Hara, B.F., Terao, A., Selby, C.P., Kilduff, T.S., Sancar, A., Edgar, D.M., and Franken, P. (2002). A role for cryptochromes in sleep regulation. *BMC Neurosci.* 3, 20.
- Zhou, J., Wang, J., Shen, B., Chen, L., Su, Y., Yang, J., Zhang, W., Tian, X., and Huang, X. (2014). Dual sgRNAs facilitate CRISPR/Cas9-mediated mouse genome targeting. *FEBS J.* 281, 1717–1725.

Cell Reports

Supplemental Information

# **Mammalian Reverse Genetics without Crossing Reveals *Nr3a* as a Short-Sleeper Gene**

Genshiro A. Sunagawa, Kenta Sumiyama, Maki Ukai-Tadenuma, Dimitri Perrin, Hiroshi Fujishima, Hideki Ukai, Osamu Nishimura, Shoi Shi, Rei-ichiro Ohno, Ryohei Narumi, Yoshihiro Shimizu, Daisuke Tone, Koji L. Ode, Shigehiro Kuraku, and Hiroki R. Ueda

# SUPPLEMENTAL EXPERIMENTAL PROCEDURES

## A Computational Model to Estimate the Minimum Efficiency to Generate Whole-body KO Mice by Multiple-target CRISPR Methods

A computational model estimated the efficiency of different CRISPR methods for generating whole-body knockout mice as follows. When sufficient Cas9 are provided, the efficiency of the single-gene mono-allelic knockout  $P$  can be described as:

$$P = \frac{S}{K+S}$$

, where total gRNA concentration and the dissociation constant between the gRNA and its target site are defined as  $S$  and  $K$ , respectively. This simple phenomenological model is adopted as the first-order approximation of the observed saturation in CRISPR-mediated DNA cutting. It describes the net effect, which is therefore intended to mechanistically include the known biochemical processes carried on during DNA cutting (e.g. PAM site binding, as well as, Cas9 remaining bonded to the cut DNA site on both strands following a cut).

When a single allele is targeted by three different gRNAs at the concentration  $S/3$  with dissociation constants ( $K_1$ ,  $K_2$ , and  $K_3$ ) between each gRNA and its target site, the single-gene mono-allelic KO efficiency  $P$  can be described as:

$$P = 1 - \left(\frac{K_1}{K_1+S/3}\right)\left(\frac{K_2}{K_2+S/3}\right)\left(\frac{K_3}{K_3+S/3}\right)$$

The concentration of each gRNA is  $S/3$  to keep total gRNA concentration as  $S$ .

Now, to include the recovery rate from single-site damage ( $\alpha$ ) is introduced as the probability to recover of a single target site from a damaged state to an intact

state. When  $\alpha$  is considered, the single-gene mono-allelic KO efficiency  $P$  for a single-target method (**Figure 1B**) can be described as:

$$P = \frac{S(1-\alpha)}{K+S}$$

Similarly, the single-gene mono-allelic KO efficiency  $P$  for a triple-target CRISPR method (**Figure 1B**) can be described as:

$$P = \frac{S(K_1K_2+K_2K_3+K_3K_1)(1-\alpha)+S^2(K_1+K_2+K_3)(1-\alpha^2)+S^3(1-\alpha^3)}{(K_1+S/3)(K_2+S/3)(K_3+S/3)}$$

When a single allele is targeted by  $N$  different sites, the dissociation constant between each gRNA and its target site is defined as  $K_i$  ( $i=1\cdots N$ ). The single-gene mono-allelic KO efficiency  $P$  for a  $N$ -tuple-target method (**Figure S1A**) can be then described as:

$$P = \frac{S(K_1\cdots K_{N-1}+\cdots+K_2\cdots K_N)(1-\alpha)+\cdots+S^{N-1}(K_1+\cdots+K_N)(1-\alpha^{N-1})+S^N(1-\alpha^N)}{\prod_{i=1}^N(K_i+S)}$$

In our simulations, the single-gene mono-allelic KO efficiency  $P$  was calculated 1,000 times for each given  $\alpha$ ,  $N$  and  $S$ . In every time of calculation, one set of  $K$  (1, 3, and  $N$  for single-, triple-, and  $N$ -tuple-target methods, respectively) were picked up randomly from a log-normal distribution whose geometric mean and geometric standard deviation as 1.0 and 2.5, respectively (**Figure S1B**). Without loss of generality, we can use 1.0 for the geometric mean of the distribution of the dissociation constant  $K$ . In order to estimate its geometric standard deviation, we first estimated three gRNA dissociation constants from the SSA assay for three gRNAs for *Tyr* gene ( $K_1 = 2.32203$ ,  $K_2 = 1.16562$  and  $K_3 = 0.369466$  for Tyr-1, Tyr-2 and Tyr-3, respectively, **Figure 1D**). We note that  $K$  is estimated as an inverse-proportional value to the DNA-cleavage efficiency measured from SSA assays. Since the geometric standard

deviation of these estimated  $K$  values was 2.53097, we used 2.5 for the geometric standard deviation. In order to estimate the efficiency of single-gene biallelic KO, double-gene biallelic KO, and triple-gene biallelic KO, we calculated  $P^2$ ,  $P^4$  and  $P^6$ , respectively when  $\alpha = 0.10$  (**Figure S1A**). Since there is uncertainty for  $\alpha$  and  $S$  *in vivo*, we also calculated single-gene biallelic KO efficiency  $P^2$  with different values of  $\alpha$  and  $S$  (**Figure S1C**,  $\alpha = 0.05, 0.10, 0.15$ , and  $S = 5, 10, 15$ , respectively). We note that calculated KO efficiencies for multiple-target methods ( $N > 1$ ) predict the “minimum” efficiency of KO since we did not include, for its simplicity, the probability of other type of mutations in the multiple-target CRISPR models (e.g. large deletions induced by multiple-target CRISPR methods), which is difficult for a DNA repair system to recover back to the intact state.

To estimate the realistic values for  $\alpha$  and  $S$ , we constructed three single-target CRISPR models with different dissociation constants ( $K_1 = 2.32203$ ,  $K_2 = 1.16562$  and  $K_3 = 0.369466$ ) estimated from the SSA assay results for Tyr-1, Tyr-2, and Tyr-3, respectively. We compared that these models with the experimental results of single-target CRISPR (36.0%, 54.2% and 64.7% for Tyr-1, Tyr-2 and Tyr-3, respectively), and used a least squares method to estimate the realistic values for a recovery rate of mutation ( $\alpha = 0.136428$ ) and total gRNA concentration ( $S = 5.7248$ ), respectively. We then performed 1,000 simulations to calculate single-gene biallelic KO efficiency ( $P^2$ ) for single-target ( $N = 1$ ) and triple-target ( $N = 3$ ) CRISPR models with  $\alpha = 0.136$  and  $K$ , which are randomly picked up from the log-normal distribution in **Figure S1B**. As a results, single-gene biallelic KO efficiency ( $P^2$ ) for single-target ( $N = 1$ ) and triple-target ( $N = 3$ ) CRISPR models at  $S = 5.7248$  are  $0.512 \pm 0.138$  and  $0.826 \pm 0.098$  (mean  $\pm$  SD), respectively (**Figure S1F**), indicating that a triple-target CRISPR method is much more efficient than a single-target CRISPR method when we used the realistic parameter values for  $\alpha$  and  $S$ . In addition, the triple-target

strategy using the mixture of Cas9 mRNA and three gRNAs achieved almost perfect efficiency (97.5%, **Figures 1E and 1F**), which is even more efficient than the predicted minimum efficiency (82.6%) (**Figure S1F**).

### **Sleep Analysis by EEG/EMG**

We anesthetized the animals and implanted telemetry transmitters (F20-EET, DSI) into them for EEG/EMG recording. The detailed methods were described previously (Sunagawa et al., 2013). The recorded data were analyzed by the FASTER method followed by a 16-second-rule filter. This rule only allows sleep or wake epochs lasting longer than 16 seconds (two epochs in this study). If the automatically staged results include sleep or wake episodes sized less than 16 seconds, these epochs will be converted to the stage just before the episode started.

### **Equivalent Electrical Circuit Simulation of the SSS System.**

Introducing the Laplace transform, the SSS system can be modeled as an electrical circuit (**Figure 3B**). Reducing the physical properties of the SSS components to those of electrical circuit elements, the pressure difference detected by the sensor can be predicted through simulation. Using these results, we optimized two variable components in the SSS system, i.e., a resistance ( $R_1$ ) and a capacitance ( $C_2$ ).

The mathematical descriptions of the model are as follows. When the total air flow is  $I$ , the flow can be divided into  $I_1$  and  $I_2$ , which are the flow to the chamber and the sensor (**Figure 3B**), respectively. How  $I$  is divided into  $I_1$  and  $I_2$  is determined

by the ratio of conductance of the chamber and the sensor unit. This conductance can be described in the Laplace domain as:

$$G_1(s) = \frac{1}{C_1 s + \frac{1}{R_1}}$$

$$G_2(s) = R_2 + \frac{1}{C_2 s}$$

where  $G_1$  and  $G_2$  denote the conductance of the SSS chamber and the pressure sensor unit, respectively.  $C_1$  is the capacitance of the SSS chamber,  $R_1$  is the resistance of the SSS chamber,  $R_2$  is the resistance of the pressure sensor, and  $C_2$  is the capacitance of the sensor. Using  $G_1$  and  $G_2$ , the air flow to the sensor can be described as:

$$I_2 = I - I_1 = \frac{G_1}{G_1 + G_2} I$$

This equation can be interpreted as  $I_2$  is a value that  $I$  is amplified by  $G_1/(G_1+G_2)$ . Now,  $R_2$  is the resistance of the sensor and  $C_1$  depends on the volume of the chamber. Using a simple flow meter and a flow pump and the pressure sensor, we estimated  $R_2$  and  $C_1$  as 152 Pa·s/ml and 0.018 ml/Pa, respectively. Because this whole circuit can be treated as a band-pass filter, our goal was to optimize  $R_1$  and  $C_2$  to maximize the gain of  $G_1/(G_1+G_2)$  over the physiological respiration frequency of mice. We ran computer simulations and selected 172 Pa·s/ml and  $6.12 \times 10^{-5}$  ml/Pa as an optimal value for  $R_1$  and  $C_2$  in our environment (**Figure 3C**).

## Fully-Automated Sleep/wake Phenotype Analysis by SSS Annotation Pipeline

The SSS annotation pipeline realizes fully automated sleep/wake phenotype analysis from respiration data of mice. The basic staging strategy was based on the FASTER method, which automatically stages the sleep state from EEG/EMG recordings (Sunagawa et al., 2013) (**Figure S2B**).

The SSS annotation pipeline is composed of three major steps (character extraction, clustering, and annotation). In the character extraction step, the recorded respiratory flow data are split into 8-second-length segments (called “epochs”). Each epoch of respiratory-flow time-domain data are next converted by fast Fourier transform (FFT) into frequency-domain data, the respiration power spectra. We used the power up to the maximum frequency, called Nyquist frequency (i.e. half of the sampling frequency), which corresponds to 125 Hz in this study because the respiratory flow was recorded at 250 Hz. Characters of respiration power spectra are then extracted by principal component analysis (PCA). The top three principal components were used in the following analysis.

In the following clustering step, we used the nonparametric density estimation clustering method (Azzalini and Torelli, 2007) to cluster the epochs. This clustering method automatically selects the number of clusters, and therefore it is powerful to cluster certain datasets without a priori model or without information on the number of clusters. A 6-day length data includes 64800 epochs. Entire epochs are divided into 12 group. Each group including 5400 epochs is used for a single-run of clustering. There are two parameters used for this clustering. One is the smoothing factor  $h_{mult}$  which is used when the probability density is estimated from the given data. The other is  $N_{grid}$ , which affects the grid density for scanning the cluster core from the estimated probability density. We used 0.80 and 270 as  $h_{mult}$  and  $N_{grid}$ , respectively.



In the final annotation step, each cluster is annotated according to the statistical value of characteristics for each cluster. During sleep the respiratory pattern represents a slow and regular flow while during wake the flow become fast and highly fluctuated (**Figure S2A**). This time-domain observation in sleep turns out to have a peaky shape in power spectrum due to the regular respiration patterns. On the other hand, during waking, the power spectrum of the respiratory flow tends to include wide range of spectra, therefore showing relatively more uniform power spectrum due to its fast and fluctuated respiration patterns. Focusing on these characteristic patterns in the frequency domain, we classified “sleep” and “awake” states by using larger and smaller skewness of the epoch’s power, respectively. “Sleep” state exhibits larger skewness than awake state because of its higher power for a narrower range of frequency and lower power for other range of frequency, which will make the distribution of sleep-epoch’s power positively skewed. Therefore, we defined when the median of skewness of epoch’s power spectrum in one cluster is within the 49th percentile (from the lowest value) of skewness of power spectrum of entire epochs, the cluster is annotated as “awake,” otherwise as “sleep”. After each cluster is annotated by skewness of power spectrum, the two epoch rule (one state continues at least two epochs) is applied to the results and the staging completes. All steps were fully-automated and unsupervised, which allows a highly-objective and reproducible sleep/wake phenotype analysis. The source code of the SSS analysis is available at **Supplemental Code S1** (<https://goo.gl/teteJM>).

The SSS technology is applied only to mice in this study. However, it might be possible to extend this method to larger organisms such as primates and human. SSS only uses a simple pressure sensor to detect the respiration pattern. Therefore, the subject needs to be placed in a semi-closed area for monitoring. Applications could be sleep monitors for babies in the incubator, or in the car for drivers.

## Calculation of Sleep/wake Parameters

Four sleep/wake parameters are mainly used in this study. One is the “sleep time.” This is the length of time that an animal’s status is staged as sleep over a given period of time. The second is the “amplitude of sleep,” or simply the “amplitude.” When an animal’s sleep time is recorded for  $D$  days and summed in the  $B$ -minutes-bin, the sleep time of the  $k^{\text{th}}$  bin is expressed as  $T_k$ , where

$$1 < k < DL_B$$

and

$$L_B = \frac{1440}{B}$$

The average sleep time of the  $m^{\text{th}}$  bin of the day is expressed as

$$\tilde{T}_m = \frac{1}{D} \sum_{i=1}^D T_{m+L_B(i-1)}$$

When the mean of all bins is expressed as

$$E_{\tilde{T}} = \frac{1}{L_B} \sum_{i=1}^{L_B} \tilde{T}_i$$

the amplitude  $A$  is defined as

$$A = \frac{1}{E_{\tilde{T}} L_B} \sum_{i=1}^{L_B} (\tilde{T}_i - E_{\tilde{T}})^2$$

The amplitude value differs with the bin “size” (i.e., duration) used to integrate the sleep time. We used a 10-min bin size in this study for amplitude calculation. Circadian and/or daily rhythmicity is expressed as a larger sleep/wake amplitude value.

The third and fourth statistics are the transition probability between sleep and awake states. In this study, sleep staging is undergone every 8 seconds. This minimal unit of sleep/wake state is called a segment (or an epoch). When we focus on two consecutive epochs, there are four possible sets of state. They are sleep to awake, awake to sleep, two sleep repeats and two wake repeats. When  $N_{mn}(X)$  is defined as a number of transitions from state  $m$  to  $n$  ( $m, n \subseteq \{\text{sleep, awake}\}$ ) within the period  $X$ ,

$$P_{ws} = \frac{N_{ws}}{N_{ws} + N_{ww}}, P_{sw} = \frac{N_{sw}}{N_{ss} + N_{sw}}.$$

The two transition probabilities,  $P_{ws}$  and  $P_{sw}$ , which are independent of each other, are indirect stabilization parameters that negatively reflect the stabilization of the awake state ( $P_{ww}$ ) and sleep state ( $P_{ss}$ ), respectively ( $P_{ws} = 1 - P_{ww}$ , and  $P_{sw} = 1 - P_{ss}$ ). We think  $P_{ws}$  and  $P_{sw}$  are useful to experimentally test the sleep models which try to provide sleep-wake temporal architecture (Diniz Behn and Booth, 2011) because the transition probabilities can be quantified by SSS.

### **Pharmacological Administration**

Methamphetamine (Philopon, Sumitomo Dainippon Pharma, Lot #006281) was diluted with natural saline to be 3 mg/kg at 100  $\mu$ l of solution. The vehicle was natural saline. Each mouse received an i.p. injection at ZT2.

Diazepam (Horizon, Maruishi Pharmaceutical, Lot #3Y140) was diluted with natural saline to be 5 mg/kg at 300  $\mu$ l of solution. The vehicle was natural saline. Each mouse received an i.p. injection at ZT13 under dark conditions.

Every animal used for pharmacological experiment was C57BL/6N male mice (6 to 8-week-old). Mice were assigned arbitrarily to control or treatment groups except the animals used in dual recording for EEG/EMG and SSS. In the dual recording, the control animals used in DZP i.p. injection were used for MAP i.p. injection after 4 days of recovery period.

### **Design of Target Sequences for gRNA**

The target sequences for *Tyr* (**Figure 1C**) were designed using Jack Lin's on-line CRISPR gRNA finder (<http://spot.colorado.edu/~slin/cas9.html>). The target sequences for *Bmal1*, *Cry1*, *Cry2*, *Per1*, *Per2*, *Hcrt* (**Figures S4A-C and S4K**), *Nr1*, *Nr2a*, *Nr2b*, *Nr2c*, *Nr2d*, *Nr3a* (set 1 and 2), and *Nr3b* (**Figures S5A-G and S6A**) were designed using the on-line CRISPR guide RNA Design tool (<http://cas9.cbi.pku.edu.cn/index.jsp>) (Ma et al., 2013). Possible off-target sequences within the mouse genome for each target sequence were checked using the CRISPR Design Tool (<http://tools.genome-engineering.org>) (Ran et al., 2013).

The alternative target sequences for *Tyr* (set 2 and 3, **Figure 7D**) were selected from the list resulting from the mm10 CRISPR/Cas9 database (<http://www.crispr.riken.jp/>).

### **Construction of the pGL3-SSA Plasmid for the Single-strand Annealing (SSA) Assay**

Two partial fragments of the pGL3-control vector (Promega), containing the 5'- or 3'-partial sequence of the *Luciferase* gene, were amplified using PCR with the following

primers: 1) forward 5'-GTAAAATCGATAAGGATCCGTCGAC-3', (Hokkaido System Science), and 2) reverse: 5'-CAGCTGAAACTGCAGAAAGATATCAAAGAATTCTTAATCCAGATCCACAACCTTCGCTTC-3' as primers amplifying part of the vector backbone and the 5' portion of the *Luciferase* gene, and 3) forward 5'-GATATCTTTCTGCAGTTTCAGCTGCCAATCATCCAAAAAATTATTATCATGG-3', and 4) reverse 5'-CATCGGTCGACGGATCCTTATCG-3' as primers amplifying part of the vector backbone and the 3' portion of the *Luciferase* gene. Both PCR products were digested by *Pst*I and *Bam*HI, and mutually ligated. The resulting vector, which contained multiple cloning sequences (5'-TAAGAATTCTTTGATATCTTTCTGCAGTTTCAGCTG-3': Stop-*Eco*RI-*Eco*RV-*Pst*I-*Pvu*II) between the 5'- and 3'-partial *Luciferase* sequences, was designated as pGL3-SSA.

### **Construction of the pSSA-Tyr-1 and pSSA-Tyr-2/3 Plasmids.**

We amplified 199-base and 200-base fragments containing the target sequences for *Tyrosinase* (Tyr-1, and both Tyr-2 and Tyr-3, respectively, **Figure S1D**) from C57BL/6 mouse genomic DNA by PCR. The PCR products were 5'-end phosphorylated with the Mighty Cloning kit (TaKaRa), and inserted into the *Eco*RV sites of the pGL3-SSA plasmid (see above). The resulting vectors were designated as pSSA-Tyr-1 and pSSA-Tyr-2/3, respectively.

## Oligonucleotide Sequences for the Target Sequences (Hokkaido System Science)

### pSSA-Tyr-1:

**Forward oligonucleotide:** 5'-GGCACCTATGGCCAAATGAACAATGGG-3'

**Reverse oligonucleotide:** 5'-GTTCCCACAATAACAAGAAAAGTCTGTGCC-3'

### pSSA-Tyr-2/3:

**Forward oligonucleotide:** 5'-TGGAACAAGCCAGTCGTATCTGGCC-3'

**Reverse oligonucleotide:** 5'-TCACAGATGGCTCTGATACAGCAAGCTG-3'

## Construction of the pX330-Tyr-1, pX330-Tyr-2, pX330-Tyr-3, pX330-Bmal1-1, pX330-Bmal1-2, and pX330-Bmal1-3 Plasmids.

Oligonucleotides (Hokkaido System Science) containing the target sequences for *Tyrosinase* and *Bmal1* (Tyr-1, Tyr-2, Tyr-3, Bmal1-1, Bmal1-2, and Bmal1-3, respectively) were annealed and inserted into the *BbsI* sites downstream of the U6 promoter on the pX330 plasmid (Cong et al., 2013; Ran et al., 2013) [Addgene #42230]. The resulting vectors were designated as pX330-Tyr-1, pX330-Tyr-2, pX330-Tyr-3, pX330-Bmal1-1, pX330-Bmal1-2, and pX330-Bmal1-3, respectively.

## Oligonucleotide Sequences for the Target Sequences

### Tyr-1:

**Forward oligonucleotide:** 5'-CACCGTGTCAAGGGACACACTGCT-3'

**Reverse oligonucleotide:** 5'-AAACAGCAGTGTGTCCCTTGACAC-3'

**Tyr-2:**

**Forward oligonucleotide:** 5'-CACCGTTATTGCTGCAGCTCTCTC-3'

**Reverse oligonucleotide:** 5'-AAACGAGAGAGCTGCAGCAATAAC-3'

**Tyr-3:**

**Forward oligonucleotide:** 5'-CACCGAAGAAGAAGCAACCCCAGG-3'

**Reverse oligonucleotide:** 5'-AAACCCTGGGGTTGCTTCTTCTTC-3'

**Bmal1-1:**

**Forward oligonucleotide:** 5'-CACCGTGTGGACTGCAATCGCAAG-3'

**Reverse oligonucleotide:** 5'-AAACCTTGCGATTGCAGTCCACAC-3'

**Bmal1-2:**

**Forward oligonucleotide:** 5'-CACCGTGTTCTGCATATTCTAACC-3'

**Reverse oligonucleotide:** 5'-AAACGGTTAGAATATGCAGAACAC-3'

**Bmal1-3:**

**Forward oligonucleotide:** 5'-CACCGTCAGATTGAAAAGAGGCGT-3'

**Reverse oligonucleotide:** 5'-AAACACGCCTCTTTTCAATCTGAC-3'

**SSA Assay**

293T cells were maintained in DMEM (Life Technologies) supplemented with 10% FBS (JRH Biosciences) and antibiotics (100 U/ml penicillin and 100 µg/ml streptomycin; Life Technologies). One day prior to transfection, the cells were plated onto six-well plates at a density of  $4 \times 10^5$  cells per well. The following day, the cells were co-transfected using FuGene6 (Roche) with 1 µg of pSSA-Tyr-1 or pSSA-Tyr-2/3 reporter plasmids in the presence of the following constructs, as

indicated in **Figure S1E**: 0 or 2 µg of pX330, pX330-Tyr-1, pX330-Tyr-2, or pX330-Tyr-3 plasmid, according to the manufacturer's instructions. Empty vector was used to bring the total amount of DNA to 3 µg per well. In addition, 50 ng of a pHRL-CMV plasmid [Renilla *luciferase* (*RLuc*) reporter vector, Promega] was included in each transfection as an internal control for transfection efficiency. Forty-eight hours after transfection, the cells were harvested and assayed using the Dual-Luciferase Reporter Assay System (Promega). The Luciferase activity was normalized to the *Rluc* activity.

### **Cas9 mRNA Synthesis**

p3s-Cas9HC plasmid(Cho et al., 2013) [Addgene #43945], which includes a T7 promoter-fused *Cas9* coding region, was digested with *Xba*I (TaKaRa), and used as the template for *in vitro* transcription using the mMACHINE T7 kit (Life Technologies). The *Cas9* mRNA was purified using the MEGAclean kit (Life Technologies).

### **gRNA Synthesis**

The gRNA templates for *Tyr* (**Figure 1C**) and *Bmal1* (**Figure S4A**) were simultaneously fused to the T7 promoter and amplified from the pX330-*Tyr* and pX330-*Bmal1* vectors, respectively (see above) by PCR using the primers listed below (Hokkaido System Science) (Wang et al., 2013). Alternatively, the gRNA templates for *Cry1*, *Cry2*, *Per1*, *Per2*, *Hcrt* (**Figures S4B, S4C and S4K**), *Nr1*, *Nr2a*, *Nr2b*, *Nr2c*, *Nr2d*, *Nr3a* (set 1 and 2), *Nr3b* (**Figures S5A-G and S6A**), and *Tyr* (**Figure 7D**) were directly synthesized and fused to the T7 promoter by PCR. First,



the partial fragments of the gRNA templates including each target sequence were amplified from pX330 plasmids (Addgene #42230) using PCR with the Common Reverse primer and Forward primer-1 (Hokkaido System Science) for each target sequence (**Table S3**). Subsequently, the T7 promoter-fused gRNA templates were amplified from the diluted PCR products using PCR with the Common Reverse primer and Forward primer-2 (Hokkaido System Science) for each target sequence (**Table S3**).

The T7 promoter-fused gRNA PCR fragments were used as the template for *in vitro* transcription using the MEGAscript T7 kit (Life Technologies). The gRNAs were purified using the MEGAclear kit (Life Technologies).

#### **Common Reverse Oligonucleotide for T7-gRNAs:**

5'-AAAAGCACCGACTCGGTGCC-3' (Wang et al., 2013)

#### **Oligonucleotide Sequences for T7-gRNAs of *Tyr* (set 1) and *Bmal1***

##### **Tyr-1 Forward oligonucleotide:**

5'-GGGCCTAATACGACTCACTATAGGTGTCAAGGGACACACTGCT-3'

##### **Tyr-2 Forward oligonucleotide:**

5'-GGGCCTAATACGACTCACTATAGGTTATTGCTGCAGCTCTCTC-3'

##### **Tyr-3 Forward oligonucleotide:**

5'-GGGCCTAATACGACTCACTATAGGAAGAAGAAGCAACCCCAGG-3'

##### **Bmal1-1 Forward oligonucleotide:**

5'-GGGCCTAATACGACTCACTATAGGTGTGGACTGCAATCGCAAG-3'

**Bmal1-2 Forward oligonucleotide:**

5'-GGGCCTAATACGACTCACTATAGGTGTTCTGCATATTCTAACC-3'

**Bmal1-3 Forward oligonucleotide:**

5'-GGGCCTAATACGACTCACTATAGGTCAGATTGAAAAGAGGCGT-3'

**Oligonucleotide Sequences for Other T7-gRNAs**

Oligonucleotide sequences of triple-target gRNAs (T7-gRNAs) for *Cry1*, *Cry2*, *Per1*, *Per2*, *Hcrt*, *Nr1*, *Nr2a*, *Nr2b*, *Nr2c*, *Nr2d*, *Nr3a* (set 1 and 2), *Nr3b*, and *Tyr* (set 2 and 3) genes are described in **Table S3**.

**Genotyping of KO Mice by Quantitative PCR (qPCR) and Sequencing**

We first prepared the genomic DNA of wild type and *Tyr* KO mice from their brains, scalps, and tails using the Wizard Genomic DNA Purification Kit (Promega, wild type, *Tyr* KO mice, **Figure S1H**) according to the manufacturer's instructions, then performed qPCR analysis for genotyping of these mice using the ABI PRISM 7900 (Applied Biosystems)/ QuantStudio7 Real-Time PCR System (Life Technologies), SYBR Premix Ex Taq GC (TaKaRa) and primers for qPCR (**Table S4**, Hokkaido System Science) and confirmed that the quantitative qPCR results matched with each other among the genomic DNAs extracted from brains, scalps and tails. Therefore, we hereafter prepared the genomic DNA only from the tails using the DNeasy Blood & Tissue Kit (QIAGEN) for genotyping of other KO mice generated for sleep/wake phenotyping (*Bmal1*, *Cry1/Cry2*, *Per1/Per2*, *Hcrt*, *Nr2a*, *Nr2b*, *Nr2c*, *Nr2d*, *Nr3a*, *Nr3b* and *Tyr* KO mice). The absolute target site abundance was calculated using a

standard curve obtained from wild-type genomic DNA. The amount of *Tbp* (Tsuji et al., 2013) was quantified and used as an internal control.

When the amplified intact DNA by qPCR is less than 0.2% of wild-type genome, we judged as the target DNA is not detectable. When either three targets of the gene were not detectable, we classified the animal as a knock-out. When all three targets were detectable, genotyping by sequencing was performed. 0.6 ~ 1 kb around the each target sequence was amplified by PCR with Ex Taq (TaKaRa) and the primers (**Table S4**, Hokkaido System Science). The fragment was subcloned into pGEM-T-easy vector (Promega), and sequenced. When every subcloned sequence has deletion and/or insertion which is not an in-frame, the target was labeled as knock-out. When the target gene has at least one knocked-out target by sequence, the animal was classified as KO. Otherwise the animal was classified as non-KO. For double-KO animals, we classified the animal as KO when both genes were knocked out by the criteria written above.

### **Protein Quantification**

Absolute amount of TYR and NR3A proteins were quantified by selected reaction monitoring (SRM) MS. Sample processing for the MS analysis was performed according to a phase-transfer surfactant (PTS) protocol (Masuda et al.) with several modifications. To analyze TYR abundance, ears from the mutant mice *Tyr* #3, *Tyr* #4, and *Nr3a* (set 2) # 10 were used. To analyze NR3A abundance, brains from the mutant mice *Tyr* #3, *Tyr* #4, *Nr3a* (set 1) # 1 - # 8, and *Nr3a* (set 2) # 1 - # 12 were used. In brief, the tissues were homogenized by sonication in PTS buffer (12 mM sodium deoxycholate, 12 mM sodium N-lauroylsarcosinate, and 50 mM  $\text{NH}_4\text{HCO}_3$ ) containing a phosphatase inhibitor cocktail (Nacalai Tesque) and a protease inhibitor

cocktail (Nacalai Tesque) and were clarified by centrifugation at 10,000xg for 10 min. The resulting homogenates were frozen in liquid nitrogen and stored at -80°C until use. Protein concentrations were determined by Quick Start Bradford Dye Reagent (Bio-Rad).

As internal standards, we added the synthetic peptides with the isotope labeled lysine and arginine residue ( $^{13}\text{C}_6$ -Lys and  $^{13}\text{C}_6$  $^{15}\text{N}_4$ -Arg, respectively). The synthetic peptides were pre-quantified and digested to produce specific sequences for the quantification of target proteins (DTLLGGSEIWR for the quantification of Tyr protein and VVTLEHPFVFTR for the quantification of NR3A protein) according to our newly developed method described elsewhere. The proteins/standards mixture were subjected to cysteine reduction and alkylation (10 mM TCEP at 37°C for 1 hr and 15 mM iodoacetamide at 37°C for 30 min in dark) followed by 5-fold dilution with 50 mM  $\text{NH}_4\text{HCO}_3$  solution. Then, enzymatic cleavage of proteins was performed by 8 hr-incubation with 1:50 (w/w) LysC followed by 16 hr-incubation with 1:50 (w/w) trypsin at 37°C. The digestion was stopped by mixing an equal volume of ethyl acetate in the presence of 0.5% TFA. The surfactants in the sample were removed by discarding the ethyl acetate phase. The left aqueous phase containing peptides was dried with SpeedVac (Thermo Scientific). The dried peptide mixture was then dissolved in an analytical buffer (2% acetonitrile and 0.1% TFA). The resulting peptide solution was desalted by using StageTip (Rappsilber et al.). Twenty  $\mu\text{g}$  of the peptide mixture prepared from brain was dissolved in 20  $\mu\text{L}$  of analytical buffer. 100  $\mu\text{g}$  of the peptide mixture prepared from ear was pre-fractionated into 6 fractions by StageTip-based fractionation (Wisniewski et al.), and then the fractions were dissolved in 14 - 30  $\mu\text{L}$  of analytical buffer.

The resulting peptide mixture corresponding to approximately 1  $\mu\text{g}$  of proteins was analyzed by liquid chromatography (LC)-MS using a triple quadrupole mass

spectrometer (TSQ Vantage EMR mass spectrometer, Thermo Scientific) equipped with a captive spray ionization source (Michrom Bioresources), a nano-Advance UHPLC system (Bruker Daltonics), and an HTC-PAL autosampler (CTC Analytics) with a trap column (0.3 x 5 mm, L-column, ODS, Chemicals Evaluation and Research Institute, Japan). Analytical samples were separated by reversed phase chromatography, in which we used a home-made capillary column (length of 200 mm and inner diameter of 100  $\mu$ m) packed with 2  $\mu$ m C18 resin (L-column2, Chemicals Evaluation and Research Institute, Japan), a gradient of acetonitrile from 4 to 36% in 0.5% acetic acid over 105 min at a flow rate 300 nL/min. Elution was directly electrosprayed (1.6 kV) into the MS.

To analyze TYR protein in ear samples, the heavy and light versions of DTLLGGSEIWR were monitored by LC-SRM-MS by using the SRM transitions of m/z 623.8 to 804.4 (for the light version) and m/z 628.8 to 814.4 (for the heavy version). To analyze NR3a protein in brain samples, the heavy and light versions of VVTLIEHPFVFTR were monitored by LC-SRM-MS by using the SRM transitions of m/z 520.0 to 766.4 (for the light version) and m/z 523.3 to 766.4 (for the heavy version).

### **Library Preparation for Exome Sequencing**

Exome libraries were constructed using SureSelectQXT Reagent Kit (Cat. No. G9681A, Agilent Technologies) and SureSelectXT Mouse All Exon Kit V1 (Cat. No. 5190-4641, Agilent Technologies) according to the manufacturer's instructions. Fragmentation and adapter-tagging of purified genomic DNA samples were performed by transposase-based reaction. After an amplification of adapter-tagged DNA library, purification was performed with AMPure XP beads (Beckman Coulter).

The purified libraries were hybridized to the SureSelect capture library, and collected hybridized DNA using streptavidin-coated beads. The enriched DNA libraries were amplified with PCR using the appropriate pair of dual indexing primers to add index tags for the Illumina TruSeq system. In each purification step, the length distribution and concentration of DNA molecules in libraries were analyzed using the 2100 Bioanalyzer (Agilent Technologies).

### **Exome Sequencing and Indel Detection**

The libraries were subjected to on-board cluster generation using TruSeq Rapid PE Cluster Kit (Cat. No. PE-402-4001) and sequenced on Rapid Run Mode of Illumina HiSeq 1500 (Illumina) to obtain paired-end reads with 126 cycles using three 50-cycle SBS kits (Cat. No. FC-402-4002), taking advantage of surplus reagents as demonstrated previously (Tatsumi K, 2015, in press.). 5% PhiX spike-in was added to each lane as control. Image analysis and base calling were executed using the standard Illumina software consisting of HiSeq Control Software (HCS) ver. 2.0.12.0 and Real-Time Analysis (RTA) ver. 1.17.21.3. Quality of the raw sequence data were controlled by FastQC ver. 0.11.1

(<http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/>). Removal of adapter sequences and low-quality reads was performed with Trim Galore ver. 0.3.3 with the parameters '-e 0.1 -q 30'

([http://www.bioinformatics.babraham.ac.uk/projects/trim\\_galore/](http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/)). Additionally, reads with the average per-base quality score of no more than 30 were removed with an original script. The reads filtered as above were aligned with the genome sequences GRCm38/mm10 by BWA ver. 0.7.10-r789 using the BWA-MEM algorithm (Li and Durbin, 2010). Potential PCR duplicates among the mapped paired reads

were removed using the MarkDuplicates function of Picard Tools ver. 1.122 (<http://picard.sourceforge.net/>). The sequencing statistics including the number of reads, mapping rate and exome on-bait coverages were summarized by the CalculateHsMetrics function of Picard Tools.

Local realignment of the mapped reads was performed by RealignerTargetCreator and IndelRealigner of the Genome Analysis Toolkit (GATK) package ver. 3.2-2 (McKenna et al., 2010), and recalibration based on per-base quality score were performed with the BaseRecalibrator and PrintReads functions. For these processes, the mm9-based exome region data with 100 bp padding provided by Agilent Technologies was converted to mm10 using the UCSC LiftOver program (<https://genome.ucsc.edu/util.html>) (Rhead et al., 2010), and insertion and deletion mutations (indels) from NCBI dbSNP (build 137) were used as known sites. The sample-specific indels were called using the HaplotypeCaller in GATK after applying local realignment and recalibration performed above.

Snpeff package ver. 4.0E (Cingolani et al., 2012) was used to filter and annotate the indels causing high-impact coding variants. We defined as high-impact mutations both the indels categorized in 'HIGH' by the Snpeff definition and the in-frame indels. The indels in both the samples and the control were removed, in comparison with the indels in the control. On- and off-target indels were visualized and evaluated on the UCSC Integrative Genomics Viewer (IGV) ver. 2.3.36 (Thorvaldsdottir et al., 2013).

The exome sequencing data have been deposited and are available in the DNA Data Bank of Japan (DDBJ). The accession numbers for *Tyr* KO and *Nr3a* KO data are DRA003478 and DRA003920 respectively.

## CRISPR gRNA Database Development

The first step in automatically detecting gRNA sites was to extract candidate targets. The process is described here for a single gene, and is repeated for all genes when running the method. Using the genome annotation file and the genomic sequence, we extracted the sequence of all exons shared by all known isoforms of the gene. This was necessary to ensure that we targeted a gene, rather than only a few of its isoforms. We then extracted all the sequences that matched the [G, C, or A]N<sub>20</sub>GG pattern and those for which the reverse-complement matched this pattern. This list represented all the possible candidates for the gene. The next steps were used to filter this list and to keep only suitable targets. Each target needed to pass all the steps, so the order did not have any impact on the selection. The steps were therefore performed in the most computationally efficient order.

The candidate gRNAs needed to target a unique site. Using Bowtie2 (Langmead and Salzberg, 2012), we eliminated all the candidates that had multiple exact matches in the genome (irrespective of location). We also removed targets with a low AT percentage (below 45%), which have a risk of binding strongly to off-target sites, and targets that contained TTTT, which tend to break the gRNA's secondary structure. Because the targets appeared in the forward primer during synthesis, we also needed to make sure they were not too similar to the reverse primer (See above "Common Reverse oligonucleotide for T7-gRNAs"). This was achieved using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970). We then used the Vienna RNA-fold package (Lorenz et al., 2011) to compute the gRNA's secondary structure. We eliminated all candidates for which the stem loop structure for Cas9 recognition could not fold (Nishimasu et al., 2014), except if the folding energy was above -18 (indicating that the 'wrong' structure was very unstable). Finally, we evaluated the off-target risk using our own implementation of the Zhang tool (Ran et



al., 2013). To ensure that all targets stored in the database were as safe as possible, we used a very strict threshold and rejected candidates with a score below 75.

Candidates that passed all the filtering steps were saved and stored in the database, which is accessible online at <http://crispr.riken.jp>.

## Statistical Analyses

Statistical analyses were performed by Microsoft Excel or R version 3.1.0(2015).

For sleep phenotype analyses of CRISPR KO mice and conventional circadian KO mice, normality was evaluated by Kolmogorov-Smirnov test at the significance level of 0.05. The homogeneity of variance for each group was evaluated by Bartlett's test when all the member of the group was normal, or Levene's test when not. In both tests, significance level was set to 0.05. The sleep phenotype of every KO strain was compared to the control (C57BL/6N, male, n=108) by Dunnett's test when the group has homogenous variance and Dunnett's Modified Tukey-Kramer pairwise multiple comparison test when not.

The histogram at the top of each sleep/wake parameter chart (**Figures 4C, F, 5C, F, and S3M**) shows the results of C57BL/6N male mice (n = 108). The error bar is the SEM for each strain, and the black dashed vertical line and gray shaded area are the mean and 1 SD range from the recordings of C57BL/6N male mice.

To estimate the sample size for phenotype evaluation, we compared the sleep time control dataset  $X_c$  (C57BL/6N, male, n=108) to a simulated dataset  $X_s$  which have the same distribution except a mean of 1 SD (50.8 min) less or much than control ( $\bar{X}_s = \bar{X}_c \pm 50.8$ ). We randomly sampled  $k$  ( $2 \leq k \leq 20$ ) data from  $X_s$  and compared with  $X_c$  by Dunnett's test with the significance level of 0.05. This test was

repeated 1000 times for one  $k$ . We calculated the rate of  $p < 0.05$  as sensitivity for each  $k$  from 2 to 20. According to this calculation, we found that the sample size  $k \geq 2$  provides more than 25% sensitivity whereas the sample size  $k \geq 6$  provides more than 50% sensitivity. In this study, we selected the sample size  $k \geq 8$  to achieve more than 75% sensitivity for the CRISPR knockout study of NMDA receptor family members.

For pharmacological studies, means of two sets of data were evaluated. First, the normality was tested by Shapiro-Wilk test at the significant level of 0.05. When the normality is confirmed in both groups, the homogeneity of variance was tested by  $F$ -test at the significance level of 0.05. When two groups were normal distributions with equal variance, Student's  $t$ -test, when the groups are normal distributions without equal variance, Welch's  $t$ -test, otherwise two-sample Wilcoxon test was applied.

For the sleep time evaluation of *Nr3a* KO mice by EEG/EMG, daily sleep time was compared between *Nr3a* KO group ( $n=2$ ) and wild type group ( $n=6$ ). The normality for wild type group was tested by Shapiro-Wilk test at the significance level of 0.05 and turned out to be normal ( $p = 0.9195$ ). Since there were only two mice tested in the *Nr3a* KO group, they were estimated to be a normal distribution because of normality confirmed in other EEG recording done in this study (the Shapiro-Wilk test for wild-type, DZP-control group, DZP-treated group, MAP-control group and MAP-treated group were 0.6609, 0.1003, 0.4282, 0.4709 and 0.3382, respectively).  $F$ -test at the significance level of 0.05 showed these groups do not have different variance ( $p = 0.3014$ ). Therefore, the mean sleep time was tested by Student's  $t$ -test ( $p = 0.0201$ , **Figure S6E**).

The i.p. injection was performed blindly without informing the injector the content of the drug. Randomization was not performed for processing data.

In this study,  $p < 0.05$  was considered as significant (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$  and n.s. for not significant evaluations).

## SUPPLEMENTAL DATA

### Figure S1. A triple-target CRISPR method for highly efficient production of whole-body biallelic KO mice, Related to Figure 1.

(A) Computational simulation predicting how efficiently one or more alleles are cut out when different numbers of gRNAs are used for the same target gene. Knockout efficiency was calculated for an increasing number of target alleles: single gene mono-allelic (one allele), single gene biallelic (two alleles), double gene biallelic (four alleles), and triple gene biallelic KO (six alleles). In each simulation, we compared single- (one target per gene), dual- (two targets), triple- (three targets), sextuple- (six targets), and decuple- (ten targets) target methods. Each method uses the same amount of total gRNA. A computational model of the different methods (except the single-target method) predicts the minimum efficiency.

(B) Distribution of the gRNA dissociation constant ( $K$ ) used in the computational simulation of CRISPR-based KO (**Figure S1A**). Distribution of the gRNA dissociation constant ( $K$ ) follows the log-normal distribution with its geometric mean and geometric standard deviation as 1.0 and 2.5, respectively. Without loss of generality, we can use 1.0 for the geometric mean in the log-normal distribution of the gRNA dissociation constant. We then estimated three gRNA dissociation constants from the SSA assay for three gRNAs for *Tyr* gene ( $K_1 = 2.32203$ ,  $K_2 = 1.16562$  and  $K_3 = 0.369466$ , **Figure 1D**). Since the geometric standard deviation of these values was 2.53097, we used 2.5 for the geometric standard deviation in the log-normal distribution of the gRNA dissociation constant.

(C) Computer simulation predicting the minimum efficiency (mean  $\pm$  SD) of single-gene biallelic KO with different recovery rates of mutations ( $\alpha = 0.05$ , 0.10 and

0.15) and total gRNA concentrations ( $S = 5, 10$  and  $15$ ). We note that total gRNA concentration ( $S$ ) is a relative value to the geometric mean of gRNA dissociation constants (1.0).

(D) Sequences of the inserted fragment for the SSA assay of gRNAs for the *Tyr* gene. The fragment sequences for the SSA vector are shown with three target sequences (blue, orange, and green) for Tyr-1, Tyr-2 and Tyr-3 gRNAs, respectively. Mouse genomic sequence data were obtained from GRCm38/mm10 via the UCSC Genome Browser (<http://genome.ucsc.edu/>) (Rhead et al., 2010).

(E) Schematic diagram of the single-strand annealing (SSA) assay. The SSA-reporter vector contained 5' and 3' *Luciferase* gene fragments that shared 702 bp of direct repeats. These fragments were separated by stop codon and gRNA target site. A gRNA/Cas9-mediated double-strand break at the target site induced an SSA reaction between the homologous regions, producing an active *Luciferase* gene.

(F) Computational simulation with estimated values for a recovery rate of mutation ( $\alpha = 0.136$ ) and total gRNA concentration ( $S = 5.72$ ) predicting how efficiently two alleles of a gene are cut out when single or triple gRNAs are used for the same target gene. We first constructed three single-target CRISPR models with three different dissociation constants ( $K_1 = 2.32203$ ,  $K_2 = 1.16562$  and  $K_3 = 0.369466$ ) estimated from the SSA assay results for Tyr-1, Tyr-2, and Tyr-3, respectively. We note that estimated values of dissociation constants ( $K_1$ ,  $K_2$  and  $K_3$ ) are inverse proportional to the DNA-cleavage efficiency in SSA assay, and geometric mean of dissociation constants is set to be 1.0 without loss of generality. The single-target CRISPR efficiencies of the experiments are shown in blue, orange and green solid circles for Tyr-1 (36.0%), Tyr-2 (54.2%) and Tyr-3 (64.7%), respectively. We then compared the single-target CRISPR models and experiments to estimate the values for a recovery

rate of mutation ( $\alpha = 0.136$ ), total gRNA concentration ( $S = 5.72$ ) by least squares method. The triple-target CRISPR efficiencies of the experiment (97.5%) and simulation (82.6%) are shown in solid and open circles (red), respectively.

(G) Coat color of the *Tyr* KO mice generated with three different single-target gRNAs. See also **Figure 1E**.

(H) The genotyping of *Tyr* KO mice. The relative amount of intact DNA for each target sequence was measured by quantitative PCR (qPCR). The genomic DNA was purified from the brain, scalp and tail of each mouse. The relative amount of intact DNA for each target sequence was scaled so that the level in one wild-type mouse was defined as 100%.

(I) An MS analysis of Tyr protein in mice skin. Mass spectrometric signals obtained by monitoring a tryptic peptide of Tyr protein (upper) in *Nr3a* (set 2) #10 (**Figure S6B**), *Tyr* #3 and #4 (**Figure S4D**) in the presence of isotopically labeled internal standard (lower) are shown. The peptide was absolutely quantified using a triple quadrupole mass spectrometer (TSQ Vantage EMR mass spectrometer, Thermo Scientific). The peptide concentration, which represents the Tyr protein concentration, was determined to be 1.7 atmol/ $\mu\text{g}$  in *Nr3a* (set 2) #10, whereas the peak was not observed in *Tyr* #3 and #4.

**Figure S2. The Snappy Sleep Stager (SSS) enables non-invasive, fully automated, high-performance sleep/wake phenotyping, Related to Figure 3.**

(A) Representative wave patterns of EEG, EMG, and the pressure change in the SSS chamber, reflecting the animal's respiration flow, during NREM sleep, REM sleep, and wake states in the mouse. During NREM sleep, the respiration rate was slow and its period was stable, whereas during waking, the respiration rate was faster and its period and amplitude fluctuated more.

(B) Schematic diagram of the SSS annotation pipeline. After the mouse is set in the chamber and the respiration flow recording started, no subjective decision is necessary until the sleep staging finishes. The recorded respiration flow is sampled at 250 Hz, then digitized, and stored in a computer. The recorded data are divided into 8-second epochs. Each epoch is transformed to a power spectrum, and its characters extracted by principal component analysis. The top three principal components are used as characters, and time-series of the characters are automatically clustered by probability density clustering. Finally, the clusters are annotated as sleep or awake based on their characteristics. See **Supplemental Experimental Procedures** for detail.

(C) Sleep/wake phenotyping using simultaneously recorded EEG/EMG and respiration. Three animals were used to estimate the SSS performance. Top row shows the principal components from the SSS analysis. Bottom row shows the time course of sleep time (per hour) determined by both EEG/EMG based staging (black line) and SSS results (red line). Note that even in different individuals, the first and second principal components clearly formed two clusters. Choosing which cluster was awake state and which was sleep state was simpler than judging all of the epoch's states in conventional analyses.

(D) Accuracy of the SSS for different levels of wakefulness. Each awake epoch was categorized as “quiet awake”, “normal awake”, or “active awake”, which were defined as 0 to 33.3%, 33.3 to 66.6%, and 66.6 to 100% EMG power, respectively. Upper panel shows representative “quiet awake” and “active awake” recordings. Lower panel shows the accuracy of the SSS results for each level of wakefulness. Note that even in “quiet awake” state the accuracy exceeded 90%.

(E) Distribution of sleep/wake parameters recorded in C57BL/6J male and female mice and C57BL/6N female mice. The upper, middle and the lower rows show the data of C57BL/6J male (n = 178), C57BL/6N female (n = 72) and C57BL/6N male (n = 173) mice, respectively. The C57BL/6J male mice slept  $699.3 \pm 49.6$  min (mean  $\pm$  SD, n = 178, 7.1% CV), which is 35 minutes less than the daily sleep time of C57BL/6N mice. In females, C57BL/6N and C57BL/6J showed  $636.4 \pm 57.7$  min and  $613.2 \pm 50.1$  min, respectively. They had shorter sleep time than males, which presumably is an effect of the estrus cycle. The pink and sky blue dashed lines show the mean and the mean  $\pm$  SD. See **Table S1**.



**Figure S3. SSS can detect various sleep/wake phenotypes, Related to Figure 4.**

(A) SSS detected the sleep-time reduction induced by a single i.p. administration of methamphetamine (MAP) at ZT2. Heatmaps show the sleep time (per hour) for all mice used in this experiment. The MAP group (n = 6) received 3 mg/kg of MAP i.p. on ZT2 of the second day, and the control group (n = 6) received vehicle i.p. at the same time. The MAP group showed less sleep than the control group for the first 3 hr after i.p. injection. Red bar at the top of the plot shows the period used for comparing the parameters in **Figure S3B**.

(B) MAP reduced the sleep time and  $P_{ws}$  significantly for the following 3 hr (ZT2 to 4) after the i.p. injection of MAP. (Left) Average sleep time. (Middle and right) Average  $P_{ws}$  and  $P_{sw}$ . Error bars represent SEM (n = 6 for each group). Welch's two sample *t*-test.

(C) A single i.p. administration of MAP reduced sleep. Plot shows the average sleep time for the MAP and control groups over time. Red bar at top indicates the 3-hr period used for comparing the groups in **Figure S3B**.

(D) SSS detects the sleep-time increase caused by a single i.p. injection of diazepam (DZP) at ZT13. Heatmaps show the sleep time (per hour) of six mice from the DZP group and the control group. The DZP group received 5 mg/kg of DZP i.p. on ZT13 of the second day, while the control group received vehicle i.p. at the same time. The DZP group showed increased sleep time for the first 5 hr after i.p. injection. Red bar at the top of the plot shows the period used for comparing the parameters in **Figure S3E**.

(E) DZP increased the sleep time and  $P_{ws}$  significantly during the 5 hr (ZT13 to 17) after i.p. administration. (Left) Total sleep time. (Middle and right) Average  $P_{ws}$  and

$P_{sw}$ . Error bars represent SEM (n = 6 for each group). Student's *t*-test or two-sample Wilcoxon test.

(F) A single i.p. administration of DZP increased sleep. Plot shows the average sleep time for the DZP and control groups. Red bar at top indicates the 5-hr period used for comparing the groups in **Figure S3E**.

(G) After DZP increased the sleep time significantly during the first 5 hr (ZT13 to 17), DZP conversely decreased the sleep time during the second 5 hr (ZT18 to 22) after i.p. administration. (Left) Total sleep time ( $p < 0.05$ ). (Middle and right) average  $P_{ws}$  and  $P_{sw}$ . Error bars represent SEM (n = 6 for each group). Student's *t*-test, Welch's *t*-test or two-sample Wilcoxon test.

(H) Performance of SSS during MAP induced sleep decrease. When 3 mg/kg of MAP i.p. was given on ZT2, the accuracy, sleep sensitivity, sleep specificity, wake sensitivity and wake specificity against EEG derived sleep status were over 90%. Error bars represent SEM (n = 4 for each group). The light and dark blue bars show the saline and MAP administered group respectively. The dashed line denotes 90%.

(I) Performance of SSS during DZP induced sleep increase. When 5 mg/kg of DZP i.p. was given on ZT13, the accuracy, sleep sensitivity, sleep specificity, wake sensitivity and wake specificity against EEG derived sleep status were over 90%. Error bars represent SEM (n = 4 for each group). The light and dark blue bars show the saline and DZP administered group respectively. The dashed line denotes 90%.

(J) Sleep time of circadian-clock mutants evaluated by SSS. Altered sleep/wake phenotypes were detected in circadian-clock mutants. Heatmaps show the sleep time (per hour) for a week under light/dark conditions for C57BL/6N (n = 108), *Bmal1* KO

mice (n = 34), *Cry1/Cry2* double-KO mice (n = 27) and *Per1/Per2* double-KO mice (n = 26), respectively.

(K) Sleep time (per hour) over 24 hr averaged over six days in circadian-clock mutants. Red lines show the mean sleep time at each time of day for the mutants. Gray line shows the data for C57BL/6N male mice (n = 108). The shaded area around the line is the SEM for each time point.

(L) Averaged  $P_{ws}$  and  $P_{sw}$  over the course of one day in circadian-clock mutants. Red lines represent the mean transition probabilities at each time of day for the mutant strains. Gray line shows the results for C57BL/6N male mice (n = 108).

(M) Distributions of sleep/wake parameters for circadian-clock mutants. Dunnett's test or Dunnett's modified Tukey-Kramer pairwise multiple comparison test compared to C57BL/6N male mice. See also **Table S1**.

**Figure S4. Combined use of a triple-target CRISPR and SSS enables sleep/wake phenotyping in single- and double-knockout mice, Related to Figure 4.**

(A-C) Target sequences of the gRNAs for knocking out clock genes. *Bmal1* (A), *Cry1* (B, left), *Cry2* (B, right), *Per1* (C, left), and *Per2* (C, right); each gene had three target sequences. Mouse genomic sequence data were obtained from GRCm38/mm10 via the UCSC Genome Browser (<http://genome.ucsc.edu/>) (Rhead et al., 2010). The colored letters (blue, orange, and green) show the 20-base target sequences. The target sequences were designed on the sense (+) or the antisense (-) strand of genomic DNA.

(D-G) The genotype of *Tyr* KO and clock KO mice. *Tyr* (D), *Bmal1* (E), *Cry1/Cry2* (f) and *Per1/Per2* (G); each gene had three targets. The relative amount of intact DNA for each target sequence was measured by quantitative PCR (qPCR). The genomic DNA was purified from the tail of each mouse. The relative amount of intact DNA for each target sequence was scaled so that the level in a wild-type mouse was defined as 100%. The numbers and wt below the bar plot denotes the mouse id and the wild-type mouse, respectively. Genotyping by sequencing was done for three *Tyr* KO mouse (D) and *Per2* gene of one *Per1/Per2* KO mice (G). The sequences of mutant alleles are shown. The target sites are colored and underlined. The mutations are labeled in red. Red texts show the description of the sequence with the number of detected allele in parentheses.

(H) Average time course of the  $P_{ws}$  and  $P_{sw}$  for *Tyr* KO and circadian-clock mutants. Red lines and red shaded area represent the mean and SEM of the transition probabilities for each time of day for the mutants. Gray line shows the results of C57BL/6N male mice (n = 108).

(I) Sleep time (per hour) phenotype under constant darkness for individual mice in circadian mutants produced by triple-target CRISPR. Sleep-time phenotype over six days is shown for the three circadian-clock mutant mice, *Bmal1* KO mice (n = 6), *Cry1/Cry2* (n = 5) and *Per1/Per2* KO mice (n = 2).

(J) Sleep time (per hour) under constant darkness over 24 hr averaged over six days in circadian mutants produced by triple-target CRISPR. Red lines and red shaded area represent the mean and SEM of sleep time at each time of day for each strain. The light gray and black bar written at the bottom of the panel denotes the subjective time of the animal.

(K) Target sequences of the gRNAs for knocking out *Hcrt* gene. Mouse genomic sequence data were obtained from GRCm38/mm10 via the UCSC Genome Browser (<http://genome.ucsc.edu/>) (Rhead et al., 2010). The colored letters (blue, orange, and green) show the 20-base target sequences. The target sequences were designed on the sense (+) or the antisense (-) strand of genomic DNA.

(L) The genotype of *Hcrt* KO mice. The relative amount of intact DNA for each target sequence was measured by qPCR. The genomic DNA was purified from the tail of each mouse. The relative amount of intact DNA for each target sequence was scaled so that the level in a wild-type mouse was defined as 100%. The numbers and wt below the bar plot denotes the mouse id and the wild-type mouse, respectively.

(M) Average time course of the  $P_{ws}$  and  $P_{sw}$  for *Hcrt* KO mice. Red lines and red shaded area represent the mean and SEM of the transition probabilities for each time of day for the mutants. Gray line shows the results of C57BL/6N male mice (n = 108).

**Figure S5. Reverse genetics without crossing revealed *Nr3a* mutant to be a short-sleeper, Related to Figure 5.**

(A-G) Target sequences of the gRNAs for the knockout of NMDA receptor family. *Nr1* (A), *Nr2a* (B), *Nr2b* (C), *Nr2c* (D), *Nr2d* (E), *Nr3a* (F), and *Nr3b* (G); each had three target sequences. The mouse genomic sequence data were obtained from GRCm38/mm10 via the UCSC Genome Browser (<http://genome.ucsc.edu/>) (Rhead et al., 2010). The colored letters (blue, orange, and green) show the 20-base target sequences. The target sequences were designed on the sense or antisense strand of genomic DNA.

(H-L) The genotype of NMDA receptor family KO mice. *Nr2a* (H), *Nr2c* (I), *Nr2d* (J), *Nr3a* (set 1) (K), and *Nr3b* (L); each had three targets. The relative amount of intact DNA for each target sequence was measured by quantitative PCR (qPCR). The genomic DNA was purified from the tail of each mouse. The relative amount of intact DNA for each target sequence was scaled so that the level in a wild-type mouse was defined as 100%. The numbers and wt below the bar plot denotes the mouse id and the wild-type mouse, respectively. Genotyping by sequencing was done for one *Nr3a* KO mouse (K) and one *Nr3b* KO mice (L). The sequences of mutant alleles are shown. The target sites are colored and underlined. The mutations are labeled in red. Red texts show the description of the sequence with the number of detected allele in parentheses.

(M) Average time course of the  $P_{ws}$  and  $P_{sw}$  of the NMDA receptor family systematic KOs for one day. Red lines and red shaded area in each chart represent the mean and SEM of the  $P_{ws}$  or  $P_{sw}$  for each time of day for the mutants. Gray line shows the results of C57BL/6N male mice (n = 108).

**Figure S6. *Nr3a* mutant is a short-sleeper in multiple KO strains and in conventional sleep analysis, Related to Figure 5.**

(A) The second set of target sequences of the gRNAs for knocking out *Nr3a* gene. Note that these sequences are distinct from set 1 sequences (**Figure S5F**). Mouse genomic sequence data were obtained from GRCm38/mm10 via the UCSC Genome Browser (<http://genome.ucsc.edu/>) (Rhead et al., 2010). The colored letters (blue, orange, and green) show the 20-base target sequences. The target sequences were designed on the sense (+) or the antisense (-) strand of genomic DNA.

(B) The genotype of *Nr3a* (set 2) KO mice. The relative amount of intact DNA for each target sequence was measured by quantitative PCR (qPCR). The genomic DNA was purified from the tail of each mouse. The relative amount of intact DNA for each target sequence was scaled so that the level in a wild-type mouse was defined as 100%. The numbers and wt below the bar plot denotes the mouse id and the wild-type mouse, respectively.

(C) Average time course of the  $P_{ws}$  and  $P_{sw}$  for *Nr3a* KO mice (set 2). Red lines and red shaded area represent the mean and SEM of the transition probabilities for each time of day for the mutants. Gray line shows the results of C57BL/6N male mice (n = 108).

(D) The genotype of *Nr3a* (set 1) KO mice used for EEG/EMG analysis. The relative amount of intact DNA for each target sequence was measured by qPCR. The genomic DNA was purified from the tail of each mouse. The relative amount of intact DNA for each target sequence was scaled so that the level in a wild-type mouse was defined as 100%. The numbers and wt below the bar plot denotes the mouse id and the wild-type mouse, respectively.

(E) The sleep time of *Nr3a* (set 1) KO evaluated by EEG/EMG based sleep staging. *Nr3a* KO group (n = 2) showed a significant short-sleeper phenotype ( $p = 0.02005$ , Student's *t*-test) than the wild-type group (n = 6). Error bars represent SEM.

(F) Mass spectrometric signals obtained by monitoring a tryptic peptide of Nr3a protein (upper) in each strain in the presence of isotopically labeled internal standard (lower) are shown. The peptide was absolutely quantified using a triple quadruple mass spectrometer (TSQ Vantage EMR mass spectrometer, Thermo Scientific). The peptide concentrations, which represents the Nr3a protein concentrations, were determined to be 2.2 atmol/ $\mu$ g in *Tyr* KO #3 and 5.7 atmol/ $\mu$ g in *Tyr* KO #4, whereas the peak was not observed in any *Nr3a* KO mice.



**Figure S7. Exome analysis of whole-body biallelic *Nr3a* (set 2) KO mice produced by triple-target CRISPR method, Related to Figure 6.**

(A) The genomic regions targeted by the three gRNAs were shown for two mice in *Nr3a* (set 2) KO group (**Figure S6B**). The upper part of each panel shows read coverage, and the lower part shows read alignments (rectangles). Horizontal lines between rectangles indicate read pairing. Different types of mutations occurring at the target sites are highlighted as follows: inter-exon deletion (red rectangles); intra-exon deletion mutation (green rectangles); short deletion (red dots); and short insertion (blue dots).

(B) Enlarged View of Read Coverage of Exome Sequence of *Nr3a* KO. Diminished depth shows a deletion mutation with the most severely affected site pointed by a red arrow. A blue vertical bar indicates an insertion mutation whose proportion is shown by its length. Percentages in parentheses are the proportions of indel mutations that can cause loss of gene function.

## SUPPLEMENTAL REFERENCES

Cho, S.W., Kim, S., Kim, J.M., and Kim, J.-S. (2013). Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nature biotechnology* 31, 230-232.

Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* 6, 80-92.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.a., *et al.* (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science (New York, NY)* 339, 819-823.

Diniz Behn, C.G., and Booth, V. (2011). Modeling the temporal architecture of rat sleep-wake behavior. *Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Conference 2011*, 4713-4716.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature methods* 9, 357-359.

Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26, 589-595.

Lorenz, R., Bernhart, S.H., Honer Zu Siederdisen, C., Tafer, H., Flamm, C., Stadler, P.F., and Hofacker, I.L. (2011). ViennaRNA Package 2.0. *Algorithms for molecular biology : AMB* 6, 26.

Masuda, T., Tomita, M., and Ishihama, Y. (2008). Phase transfer surfactant-aided trypsin digestion for membrane proteome analysis. *J Proteome Res* 7, 731-740.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytzky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., *et al.* (2010). The Genome Analysis

Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research* 20, 1297-1303.

Needleman, S.B., and Wunsch, C.D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of molecular biology* 48, 443-453.

Nishimasu, H., Ran, F.A., Hsu, P.D., Konermann, S., Shehata, S.I., Dohmae, N., Ishitani, R., Zhang, F., and Nureki, O. (2014). Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* 156, 935-949.

R core team (2015). R: A Language and Environment for Statistical Computing, R.C. Team, ed. (Vienna, Austria: R Foundation for Statistical Computing).

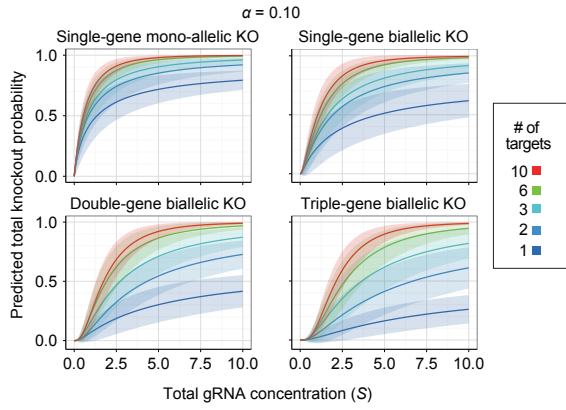
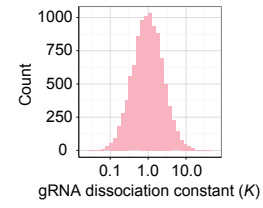
Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature protocols* 2, 1896-1906.

Tatsumi K, N.O., Itomi K, Tanegashima C, Kuraku S (2015, in press.). Optimization and cost-saving in tagmentation-based mate-pair library preparation and sequencing. *Biotechniques*.

Thorvaldsdottir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in bioinformatics* 14, 178-192.

Tsujino, K., Narumi, R., Masumoto, K.H., Susaki, E.A., Shinohara, Y., Abe, T., Iigo, M., Wada, A., Nagano, M., Shigeyoshi, Y., *et al.* (2013). Establishment of TSH beta real-time monitoring system in mammalian photoperiodism. *Genes to cells : devoted to molecular & cellular mechanisms* 18, 575-588.

Wisniewski, J.R., Zougman, A., and Mann, M. (2009). Combination of FASP and StageTip-based fractionation allows in-depth analysis of the hippocampal membrane proteome. *J Proteome Res* 8, 5674-5678.

**A****B****C**

Single-gene biallelic KO efficiency			
	Total gRNA concentration (S)		
	5	10	15
Recovery rate of mutation ( $\alpha$ )	0.05	0.10	0.15
	0.85 $\pm$ 0.11	0.95 $\pm$ 0.05	0.97 $\pm$ 0.03
	0.83 $\pm$ 0.11	0.93 $\pm$ 0.06	0.96 $\pm$ 0.03
	0.79 $\pm$ 0.12	0.90 $\pm$ 0.07	0.93 $\pm$ 0.04

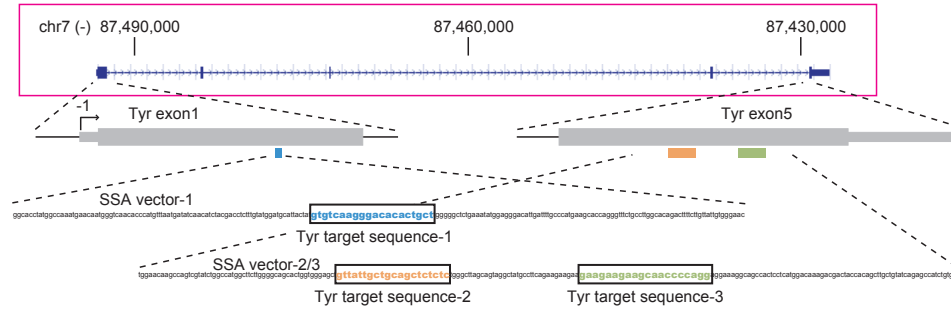
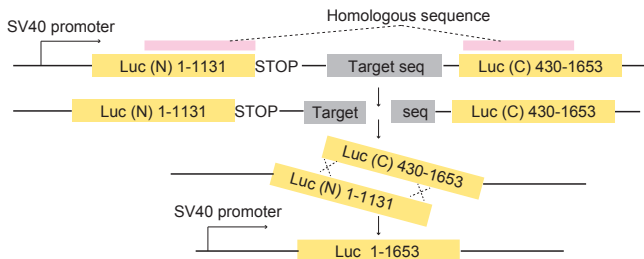
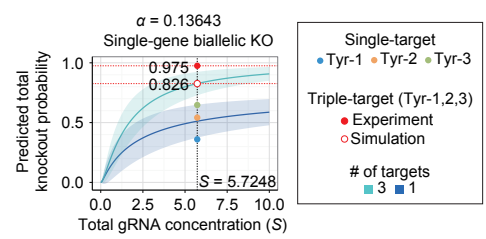
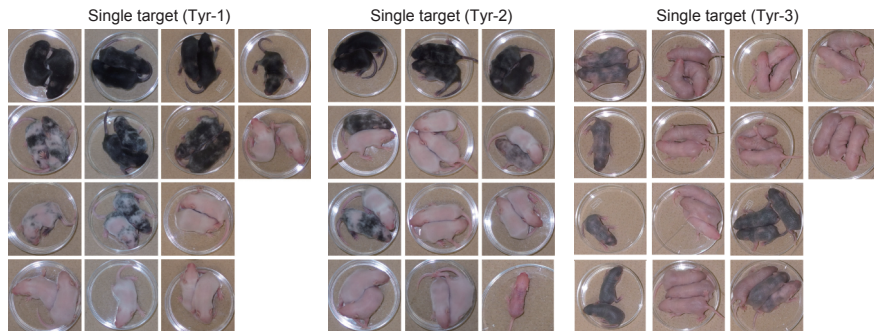
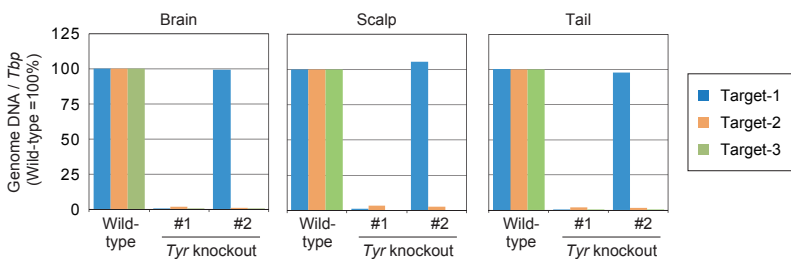
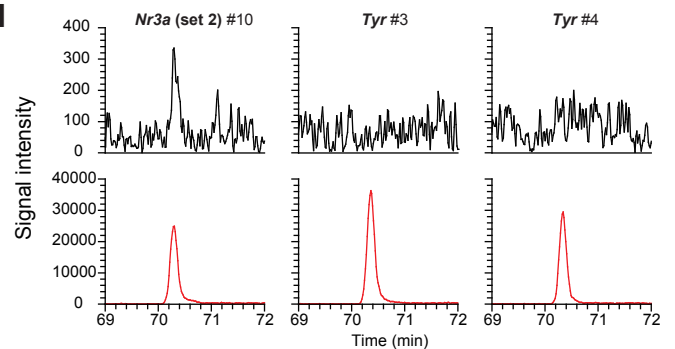
**D****E****F****G****H****I**

Figure S1 G.A. Sunagawa

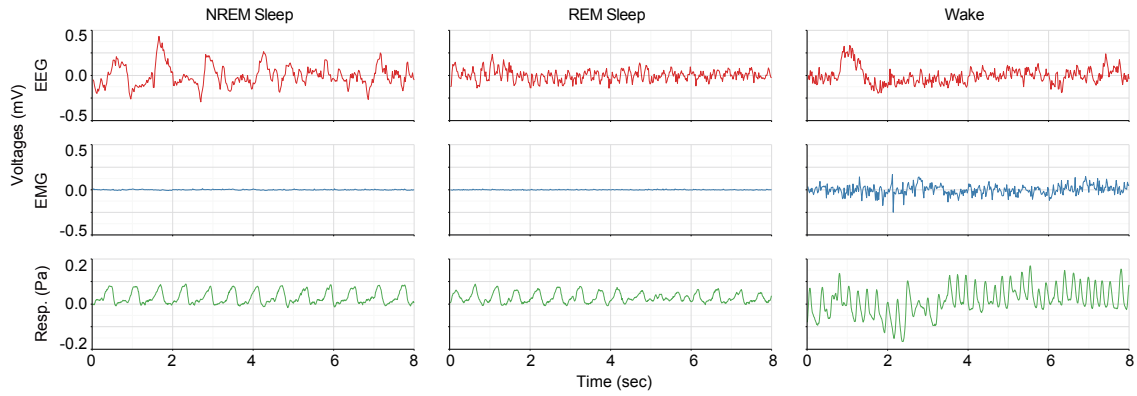
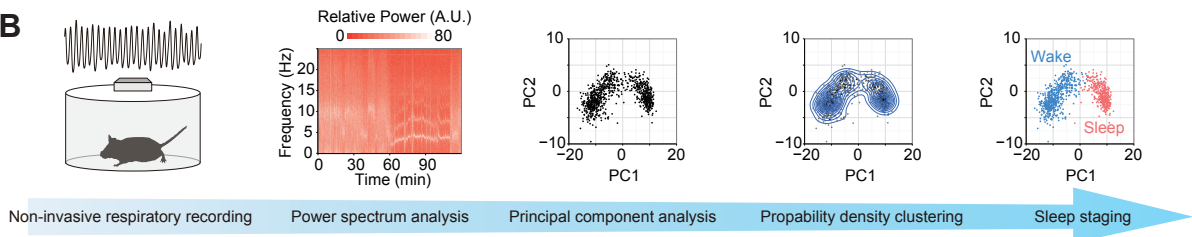
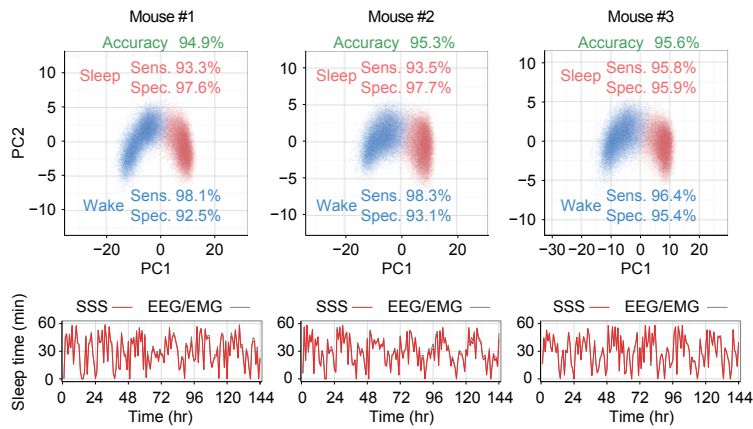
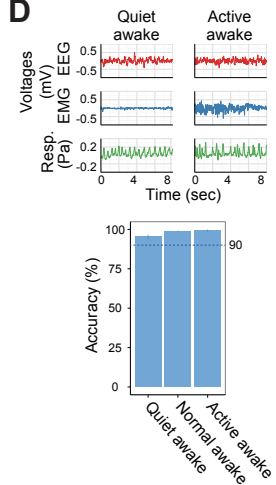
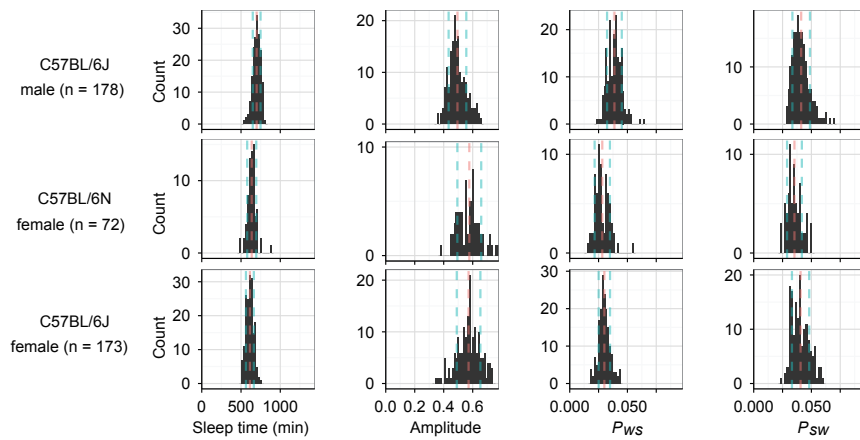
**A****B****C****D****E**

Figure S2 G.A. Sunagawa

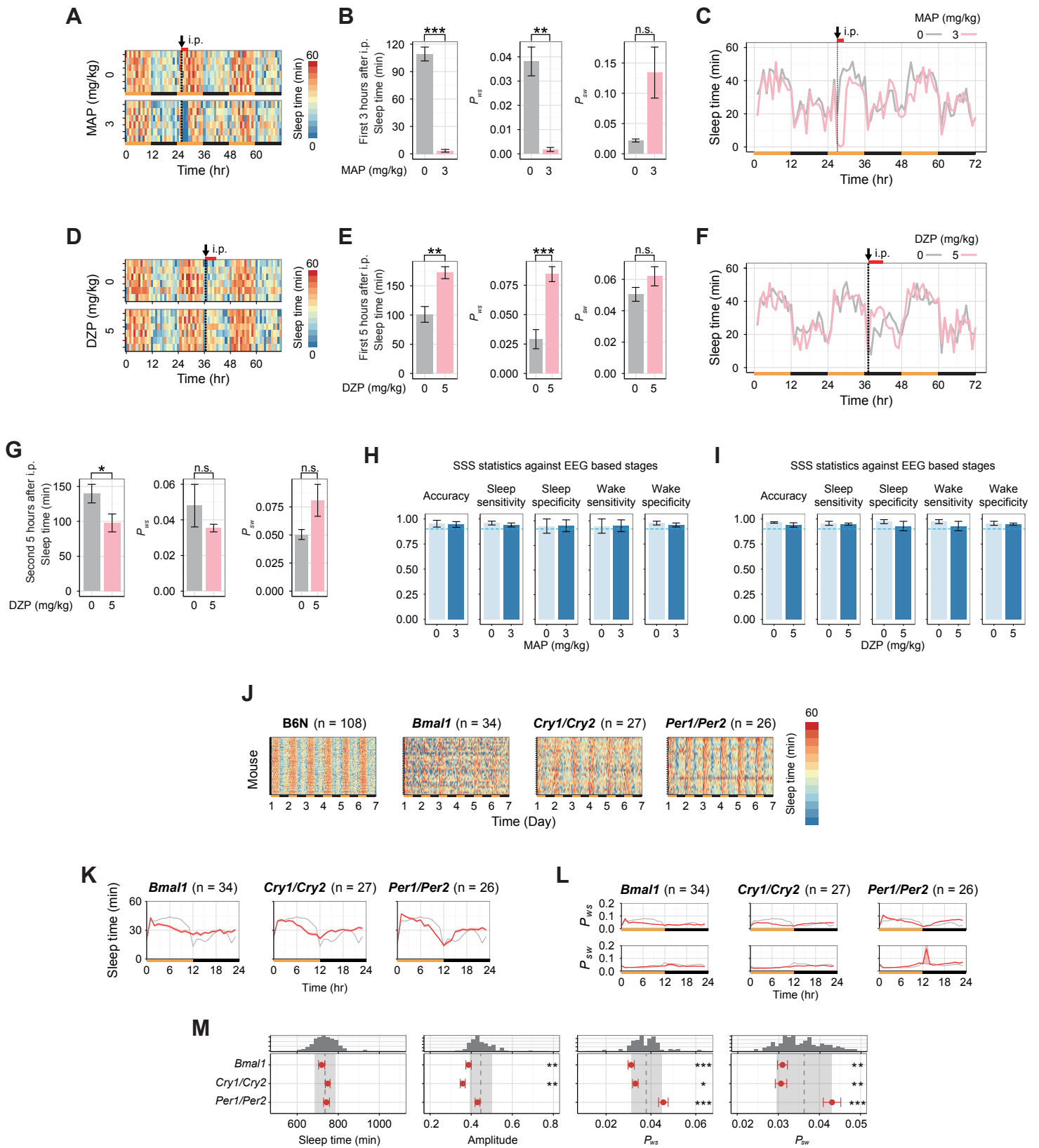


Figure S3 G.A. Sunagawa

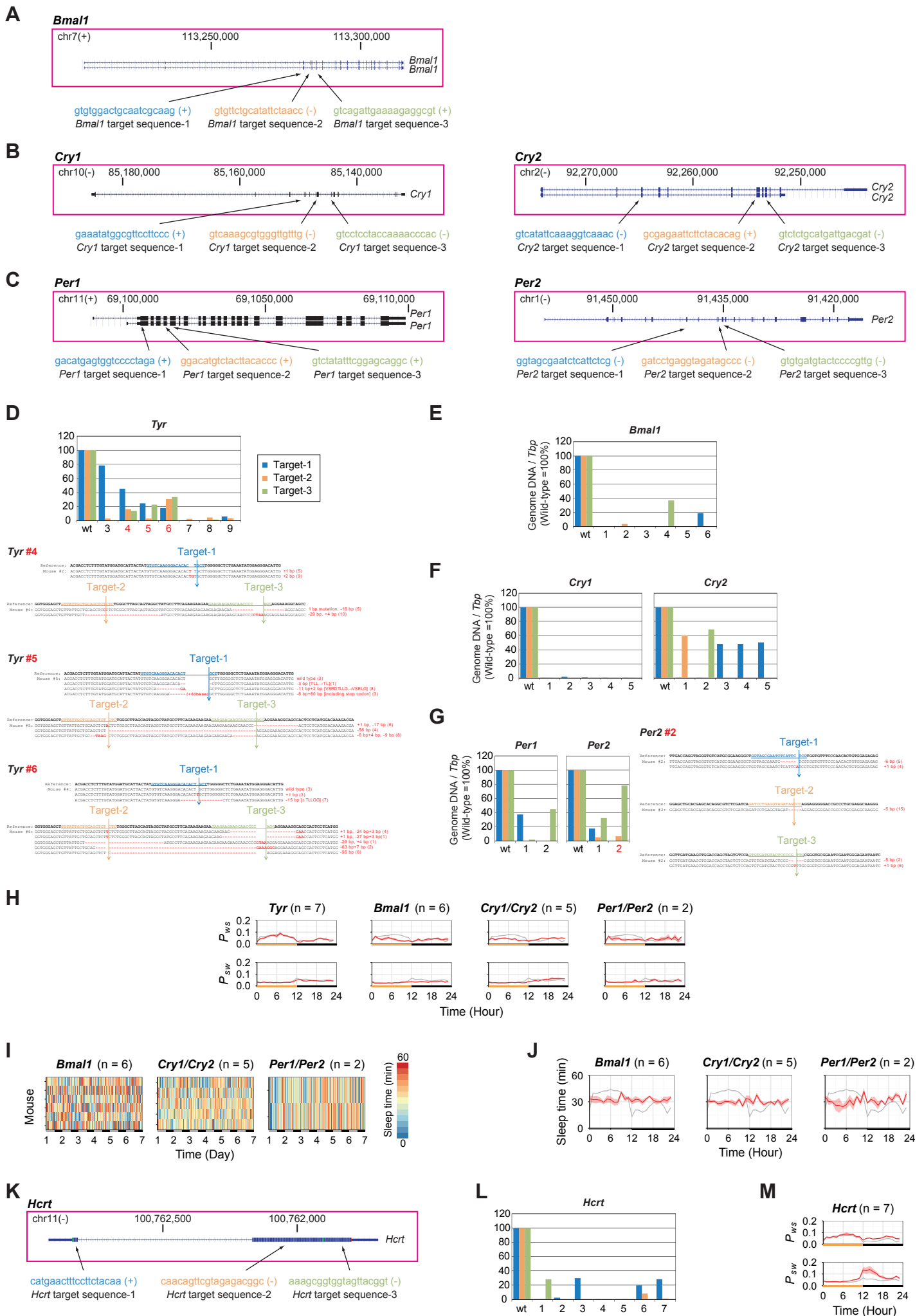


Figure S4 G.A. Sunagawa

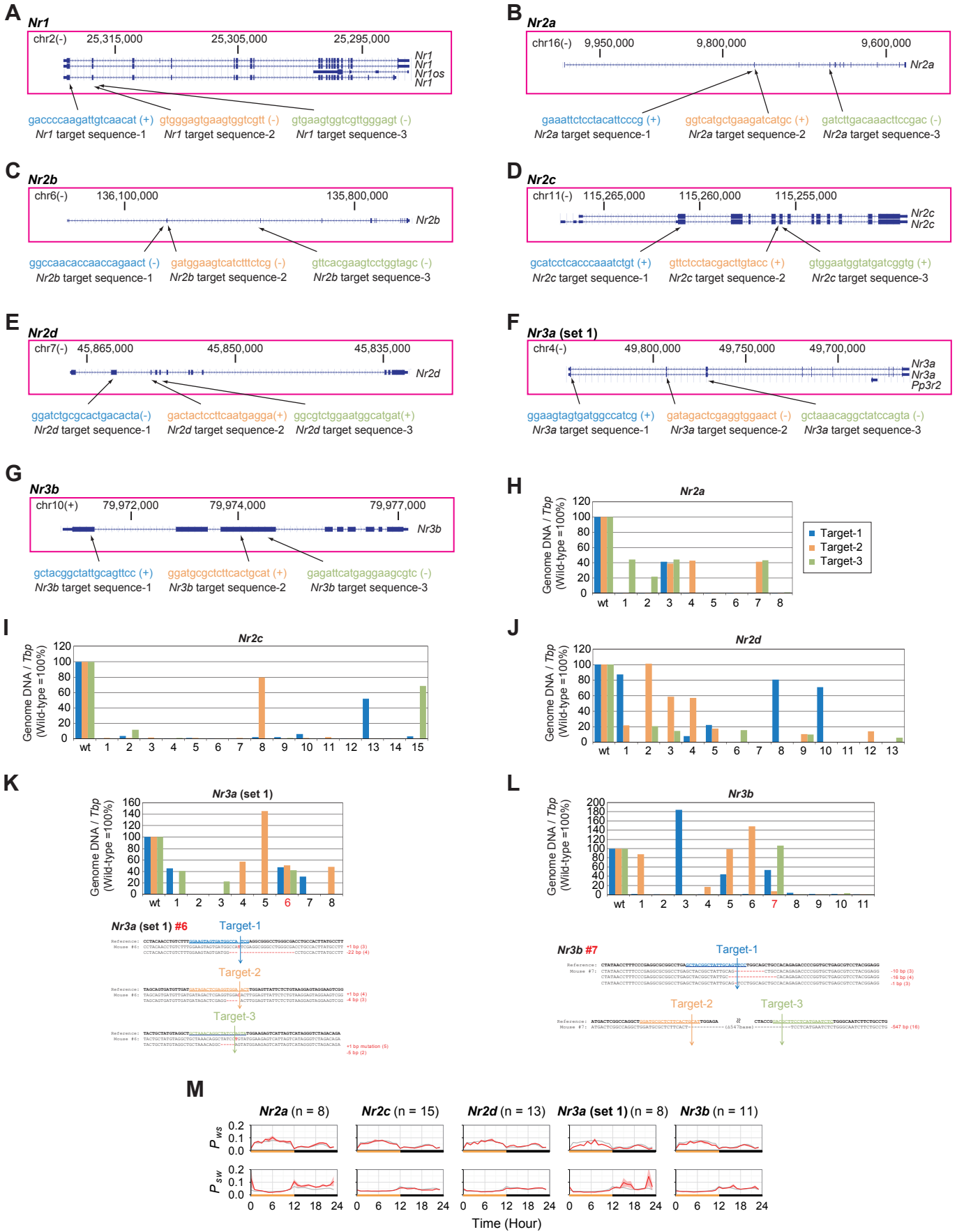


Figure S5 G.A. Sunagawa



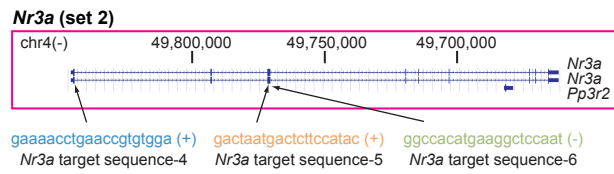
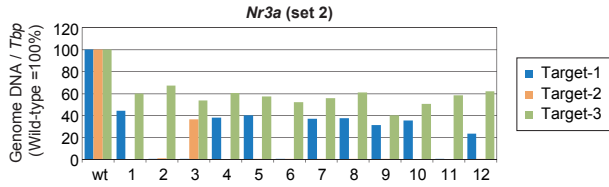
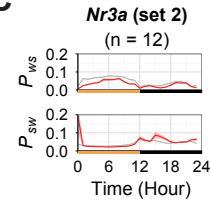
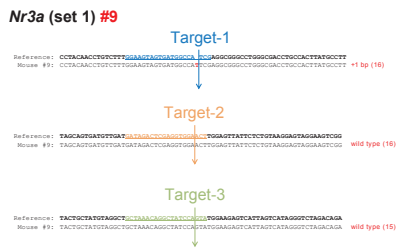
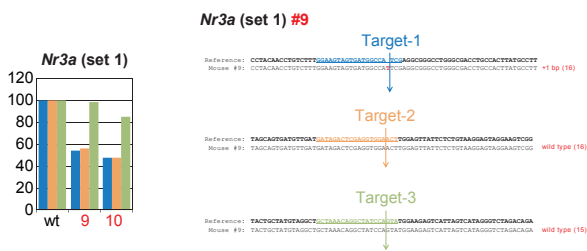
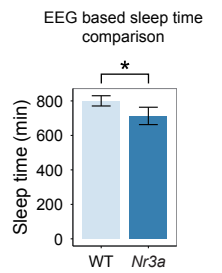
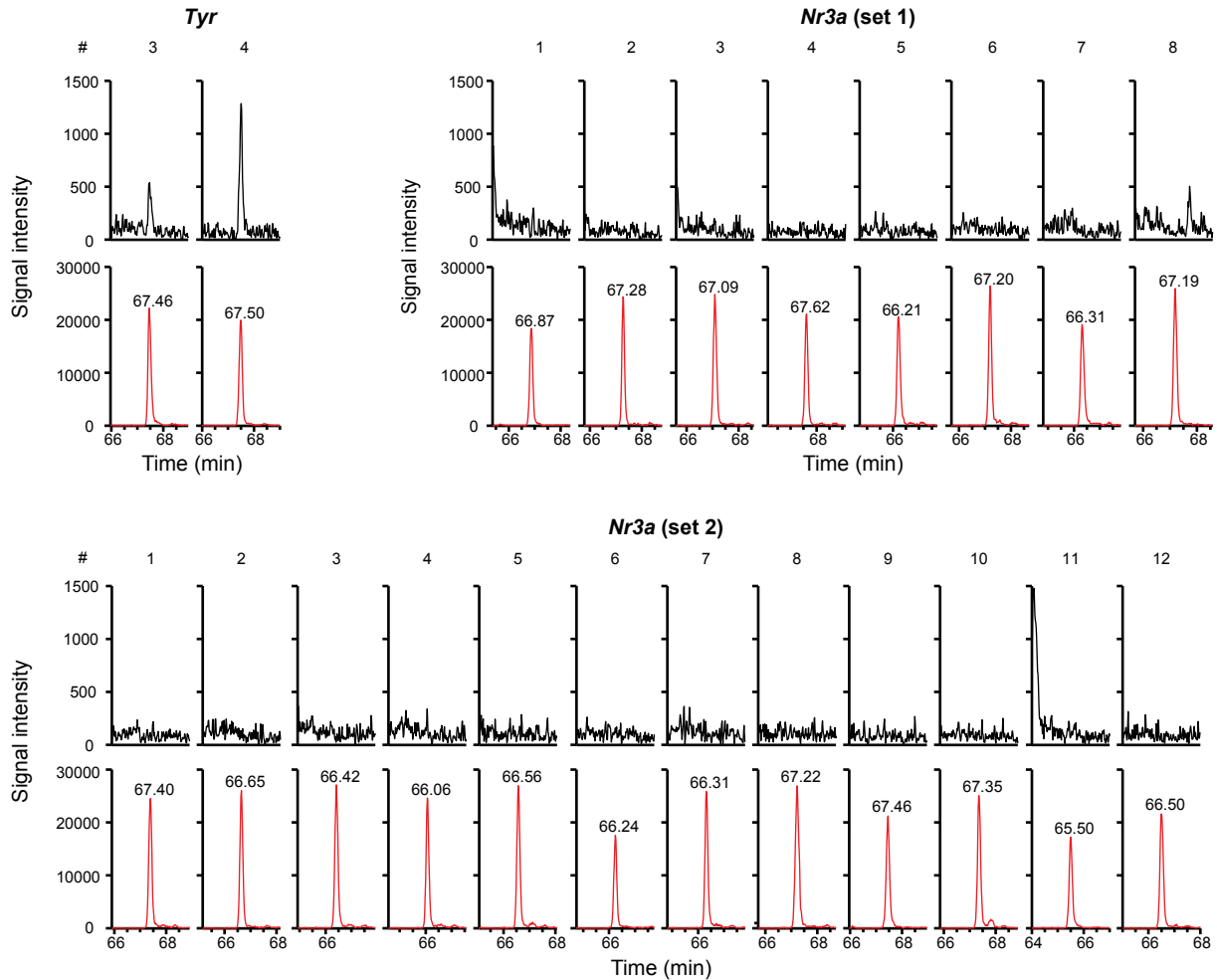
**A****B****C****D****E****F**

Figure S6 G.A. Sunagawa

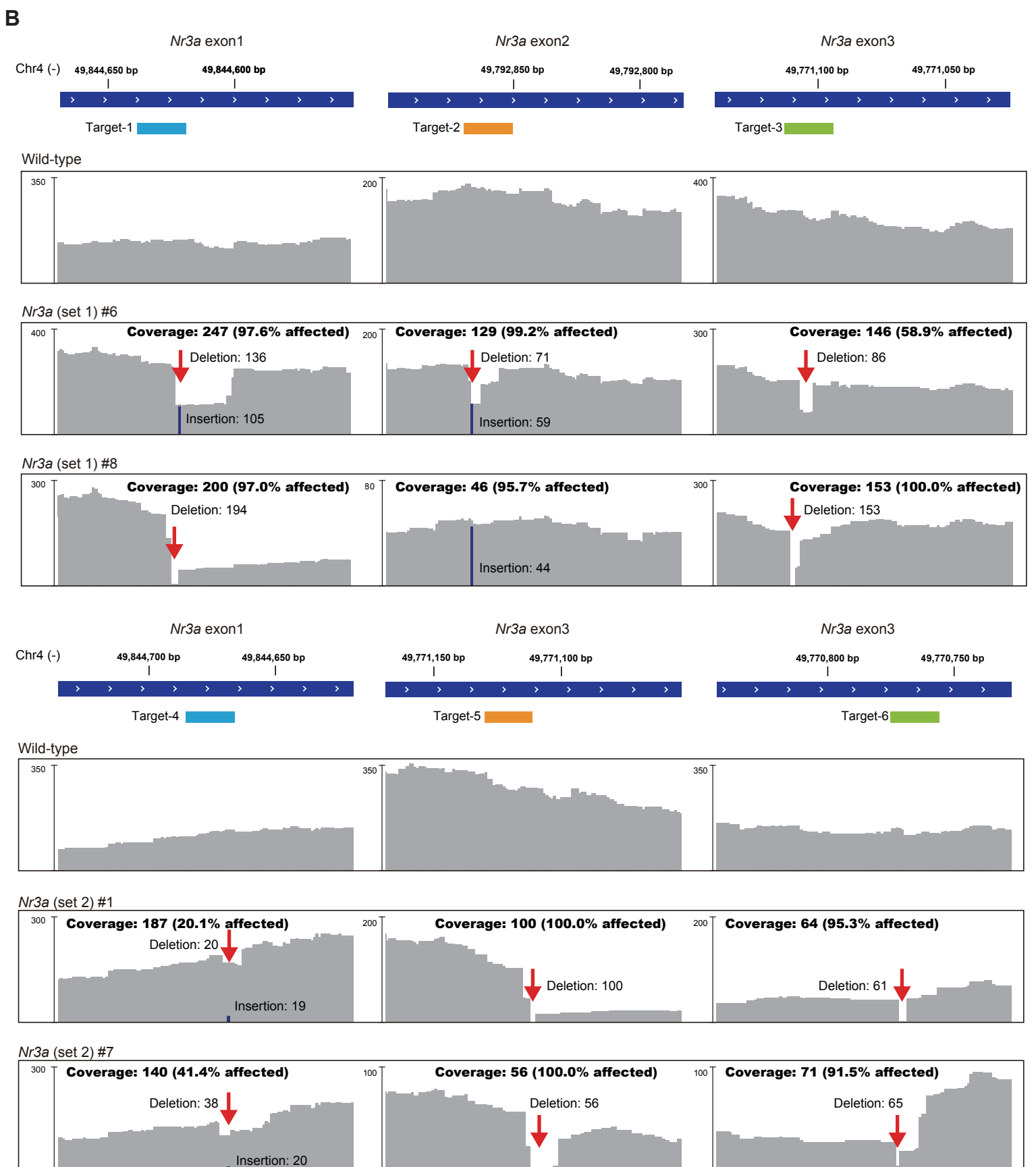
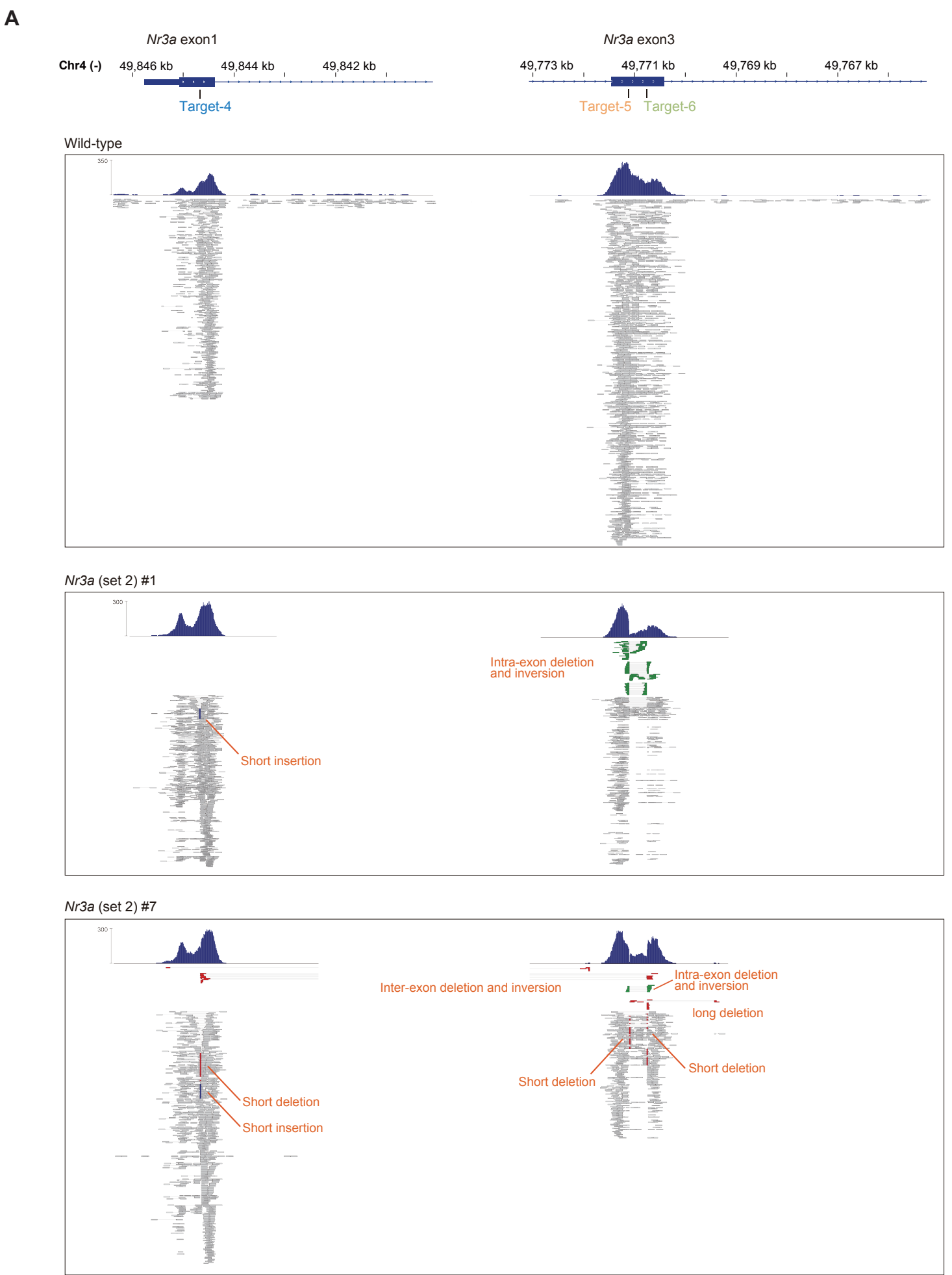


Figure S7 G.A. Sunagawa

**Table S1. Sleep/wake phenotype, Related to Figure 3, 4 and 5.** Every mouse phenotyped in SSS under LD conditions. The values are mean  $\pm$  SD for wild-type and mean  $\pm$  SEM for KOs.

	Strain/gene	#	Daily sleep time (min)	Amplitude	$P_{ws}$	$P_{sw}$
Wild-type	C57BL/6N male	108	735.7 $\pm$ 50.8	0.4462 $\pm$ 0.0530	0.0379 $\pm$ 0.0068	0.0363 $\pm$ 0.0067
	C57BL/6N female	72	636.4 $\pm$ 57.7	0.5764 $\pm$ 0.0823	0.0282 $\pm$ 0.0066	0.0353 $\pm$ 0.0064
	C57BL/6J male	178	699.3 $\pm$ 49.6	0.4948 $\pm$ 0.0610	0.0387 $\pm$ 0.0064	0.0411 $\pm$ 0.0076
	C57BL/6J female	173	613.2 $\pm$ 50.1	0.5716 $\pm$ 0.0815	0.0300 $\pm$ 0.0050	0.0406 $\pm$ 0.0075
Conventional KO	<i>Bmal1</i>	34	719.1 $\pm$ 15.0	0.3871 $\pm$ 0.0132	0.0312 $\pm$ 0.0014	0.0310 $\pm$ 0.0012
	<i>Cry1/Cry2</i>	27	748.8 $\pm$ 11.5	0.3586 $\pm$ 0.0119	0.0331 $\pm$ 0.0012	0.0307 $\pm$ 0.0014
	<i>Per1/Per2</i>	26	742.9 $\pm$ 14.7	0.4308 $\pm$ 0.0119	0.0456 $\pm$ 0.0021	0.0431 $\pm$ 0.0021
Triple-target CRISPR KO	<i>Tyr</i>	7	734.5 $\pm$ 17.2	0.4249 $\pm$ 0.0097	0.0368 $\pm$ 0.0023	0.0350 $\pm$ 0.0012
	<i>Bmal1</i>	6	758.5 $\pm$ 14.6	0.3735 $\pm$ 0.0189	0.0307 $\pm$ 0.0022	0.0274 $\pm$ 0.0015
	<i>Cry1/Cry2</i>	5	725.8 $\pm$ 14.5	0.3489 $\pm$ 0.0092	0.0388 $\pm$ 0.0018	0.0381 $\pm$ 0.0013
	<i>Per1/Per2</i>	2	716.2 $\pm$ 2.3	0.3699 $\pm$ 0.0253	0.0364 $\pm$ 0.0009	0.0368 $\pm$ 0.0007
	<i>Hcrt</i>	7	720.1 $\pm$ 24.1	0.3925 $\pm$ 0.0278	0.0548 $\pm$ 0.0022	0.0556 $\pm$ 0.0048
	<i>Nr2a</i>	8	711.9 $\pm$ 29.5	0.5328 $\pm$ 0.0290	0.0407 $\pm$ 0.0035	0.0410 $\pm$ 0.0026
	<i>Nr2c</i>	15	715.0 $\pm$ 11.6	0.4557 $\pm$ 0.0188	0.0343 $\pm$ 0.0012	0.0348 $\pm$ 0.0013
	<i>Nr2d</i>	13	755.8 $\pm$ 10.6	0.4052 $\pm$ 0.0131	0.0408 $\pm$ 0.0011	0.0369 $\pm$ 0.0010
	<i>Nr3a</i> (set 1)	8	639.0 $\pm$ 17.5	0.5826 $\pm$ 0.0342	0.0250 $\pm$ 0.0015	0.0310 $\pm$ 0.0009
	<i>Nr3a</i> (set 2)	12	627.1 $\pm$ 11.2	0.5737 $\pm$ 0.0169	0.0254 $\pm$ 0.0009	0.0363 $\pm$ 0.0006
	<i>Nr3b</i>	11	683.7 $\pm$ 11.3	0.5004 $\pm$ 0.0149	0.0310 $\pm$ 0.0015	0.0342 $\pm$ 0.0015

**Table S2. KO mice production by triple-target CRISPR method, Related to Figure 1, 4 and 5.** Cas9 mRNA and gRNAs targeting *Tyr* gene and circadian/sleep related genes (*Bmal1*, *Cry1/Cry2*, *Per1/Per2* and *Hcrt*) and NMDA receptor family (*Nr1*, *Nr2a*, *Nr2b*, *Nr2c*, *Nr2d*, *Nr3a*, and *Nr3b*) were injected into fertilized C57BL/6N eggs. Male mice, which survived at least the age of 6 weeks, were used for SSS analysis. The table shows the number of transferred embryos, number of offsprings and the results of genotyping for all mice used in SSS

Gene	Transferred embryos	Offspring		Genotyped male mice	KO confirmed male mice	Non-KO confirmed male mice	Not-determined male mice
		Male	Female				
<i>Tyr</i>	150	9	11	7	7	0	0
<i>Bmal1</i>	111	6	8	6	6	0	0
<i>Cry1/Cry2</i>	105	6	8	6	5	0	1
<i>Per1/Per2</i>	169	3	5	3	2	0	1
<i>Hcrt</i>	96	11	8	10	7	0	3
<i>Nr1</i>	205	0	0	0	0	0	0
<i>Nr2a</i>	131	8	12	8	8	0	0
<i>Nr2b</i>	236	1	0	1	0	1	0
<i>Nr2c</i>	143	16	8	15	15	0	0
<i>Nr2d</i>	140	13	20	13	13	0	0
<i>Nr3a</i> (set 1)	135	8	9	8	8	0	0
<i>Nr3a</i> (set 2)	124	13	7	13	12	0	1
<i>Nr3b</i>	205	12	9	12	11	0	1
Total	1950	106	105	102	94	1	7
				100.0%	92.2%	1.0%	6.9%

**Table S5. Exome sequencing statistics for *Tyr* and *Nr3a* KO mice, Related to Figure 2 and 6.** See Methods for the detail of library preparation and sequencing. 'Target bases' are those in bait regions designed for capturing exons.

Library	Total number of reads	Number of uniquely mapped reads	Uniquely mapped reads (%)	Number of on-bait bases	Proportion of on-bait bases in uniquely mapped reads (%)	Target bases achieving 10X or greater coverage (%)	Sample name in DDBJ
Wildtype used in <i>Tyr</i> KO analysis	50,520,378	46,336,283	91.7	3,032,197,583	53.7	95.0	tyr-wild-type
<i>Tyr</i> #1	53,635,574	49,124,972	91.6	3,196,839,130	53.5	95.6	tyr-01
<i>Tyr</i> #2	50,975,220	47,102,814	92.4	2,258,383,004	39.4	91.6	tyr-02
Wildtype used in <i>Nr3a</i> KO analysis	237,830,634	217,074,710	91.3	13,453,459,028	51.0	99.0	nr3a-wild-type
<i>Nr3a</i> (set 1) #6	182,374,702	172,485,979	94.6	14,927,922,341	69.9	98.9	nr3a-set1-06
<i>Nr3a</i> (set 1) #8	194,956,048	184,558,103	94.7	15,510,369,211	67.8	99.0	nr3a-set1-08
<i>Nr3a</i> (set 2) #1	175,181,378	167,133,655	95.4	14,427,136,154	69.7	98.5	nr3a-set2-01
<i>Nr3a</i> (set 2) #7	187,343,854	177,780,377	94.9	15,194,075,796	69.0	98.8	nr3a-set2-07

Data S1. The Source Code of the SSS Analysis, Written in R Language

The code can be downloaded from <https://goo.gl/teteJM>.