doi:10.1038/nature16533

Autism-like behaviours and germline transmission in transgenic monkeys overexpressing MeCP2

Zhen Liu[1]*, Xiao Li[1]*, Jun-Tao Zhang[1], Yi-Jun Cai[1], Tian-Lin Cheng[1], Cheng Cheng[1], Yan Wang[1], Chen-Chen Zhang[1], Yan-Hong Nie[1], Zhi-Fang Chen[1], Wen-Jie Bian[1], Ling Zhang[2], Jianqiu Xiao[2], Bin Lu[1], Yue-Fang Zhang[1], Xiao-Di Zhang[1], Xiao Sang[1], Jia-Jia Wu[1], Xiu Xu[3], Zhi-Qi Xiong[1], Feng Zhang[2], Xiang Yu[1], Neng Gong[1], Wen-Hao Zhou[4], Qiang Sun[1] & Zilong Qiu[1]

Methyl-CpG binding protein 2 (MeCP2) has crucial roles in transcriptional regulation and microRNA processing[1–4]. Mutations in the *MECP2* gene are found in 90% of patients with Rett syndrome, a severe developmental disorder with autistic phenotypes[5]. Duplications of *MECP2*-containing genomic segments cause the *MECP2* duplication syndrome, which shares core symptoms with autism spectrum disorders[6]. Although *Mecp2*-null mice recapitulate most developmental and behavioural defects seen in patients with Rett syndrome, it has been difficult to identify autism-like behaviours in the mouse model of MeCP2 overexpression[7,8]. Here we report that lentivirus-based transgenic cynomolgus monkeys (*Macaca fascicularis*) expressing human MeCP2 in the brain exhibit autism-like behaviours and show germline transmission of the transgene. Expression of the *MECP2* transgene was confirmed by western blotting and immunostaining of brain tissues of transgenic monkeys. Genomic integration sites of the transgenes were characterized by a deep-sequencing-based method. As compared to wild-type monkeys, *MECP2* transgenic monkeys exhibited a higher frequency of repetitive circular locomotion and increased stress responses, as measured by the threat-related anxiety and defensive test[9]. The transgenic monkeys showed less interaction with wild-type monkeys within the same group, and also a reduced interaction time when paired with other transgenic monkeys in social interaction tests. The cognitive functions of the transgenic monkeys were largely normal in the Wisconsin general test apparatus, although some showed signs of stereotypic cognitive behaviours. Notably, we succeeded in generating five F_1 offspring of *MECP2* transgenic monkeys by intracytoplasmic sperm injection with sperm from one F_0 transgenic monkey, showing germline transmission and Mendelian segregation of several *MECP2* transgenes in the F_1 progeny. Moreover, F_1 transgenic monkeys also showed reduced social interactions when tested in pairs, as compared to wild-type monkeys of similar age. Together, these results indicate the feasibility and reliability of using genetically engineered non-human primates to study brain disorders.

We first co-injected lentivirus expressing synapsin-promoter-driven[10] haemagglutinin (HA)-tagged human MeCP2 and green fluorescence protein (GFP) and lentivirus expressing mCherry into the perivitelline space of 94 mature oocytes of cynomolgus monkeys (Fig. 1a). We found that 61 out of 88 (69%) of the surviving oocytes became zygotes after intracytoplasmic sperm injection (ICSI), and 53 embryos were then transferred into 18 surrogate monkeys. Nine surrogates (9 out of 18, 50%) became pregnant and produced eight live births (3 male, 5 female; Fig. 1b) and four stillbirths, all carrying human *MECP2*, GFP and mCherry transgenes, as determined by PCR (Fig. 1c). The AccuCopy assay showed that the copy numbers of *MECP2* transgenes in 8 live (T04–T11) and 2 aborted (T01 and T02) transgenic (TG) monkeys varied from 1.0 to 7.3 (Extended Data Table 1a). In the second experiment, we injected 264 mature oocytes with lentivirus carrying the hSynapsin-HA-hMECP2-2a-GFP cassette, and transferred 105 embryos after ICSI into 36 surrogates. Owing to unfavourable seasonal conditions, only 7 pregnant surrogates gave birth to 9 monkeys (T13–T21), and only 2 survived (Supplementary Table 1).

Western blotting of tissues of stillbirth TG monkey T14 showed expression of GFP and HA–MeCP2 proteins in the cortex and cerebellum, but not in non-neural tissues, confirming specific transgene expression under the synapsin promoter (Fig. 1d). Levels of MeCP2 protein were also significantly higher than that found in an aborted wild-type (WT) monkey of a similar age (Fig. 1e, f). Transgenic integration was confirmed by Southern blotting using a probe targeting the HA-hMECP2-2a-GFP transgene (Extended Data Fig. 1a). We next analysed genomic integration sites of lentiviral cassettes containing HA-hMECP2-2a-GFP and mCherry transgenes by a deep-sequencing-based method (Extended Data Fig. 1b). All transgenes were located in genomic loci distant from known coding exons, and thus unlikely to interfere with endogenous genes (Fig. 1g and Supplementary Table 2), and insertion numbers were largely consistent with the copy numbers identified by AccuCopy (Extended Data Fig. 1c). Therefore, the human *MECP2* transgene was successfully incorporated into the monkey genome and specifically expressed in the monkey's brain.

Compared to WT monkeys of similar ages, the body weight and abdominal circumference of the TG group (T04–T11) was slightly lower before 20 months of age, whereas no difference was found for head–truck length, heart and respiratory rates or body temperature (Extended Data Fig. 2a–g). We did not observe in TG monkeys any seizure phenotype, which was associated with *MECP2* duplication syndrome patients[6], perhaps owing to the young age of the monkeys. Interestingly, mass spectrometry of blood metabolites at ~18 and ~36 months suggested metabolic abnormalities in the TG group, with significantly higher levels of some short- and long-chain fatty acids (Extended Data Fig. 3a, b), reminiscent of some human autistic patients[11].

Despite their generally normal early development, one TG monkey (T05) showed severe weight loss and head circumference reduction after 15 months (Extended Data Fig. 4a–c), and was unable to complete behavioural tests. Monkeys T09 and T07 became severely sick at 43 and 46 months, respectively, after behavioural tests. The sickness of these TG monkeys echoed the severe phenotypes of human patients with the *MECP2* duplication syndrome[6]. The euthanasia procedure was performed, and their brain tissues were collected for further analysis with western blotting, immunostaining and RNA-sequencing (RNA-seq). We found that HA–MeCP2 and GFP were expressed in the brain

[1]Institute of Neuroscience, CAS Key Laboratory of Primate Neurobiology, State Key Laboratory of Neuroscience, CAS Center for Excellence in Brain Science and Intelligence Technology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 320 Yue-Yang Road, Shanghai 200031, China. [2]State Key Laboratory of Genetic Engineering and Ministry of Education Key Laboratory of Contemporary Anthropology, Collaborative Innovation Center of Genetics and Development, School of Life Sciences, Fudan University, Shanghai 200438, China. [3]Department of Child Healthcare, Children's Hospital of Fudan University, Shanghai 201102, China. [4]Department of Neonatology, Children's Hospital of Fudan University, Shanghai 201102, China.
*These authors contributed equally to this work.

© 2016 Macmillan Publishers Limited. All rights reserved

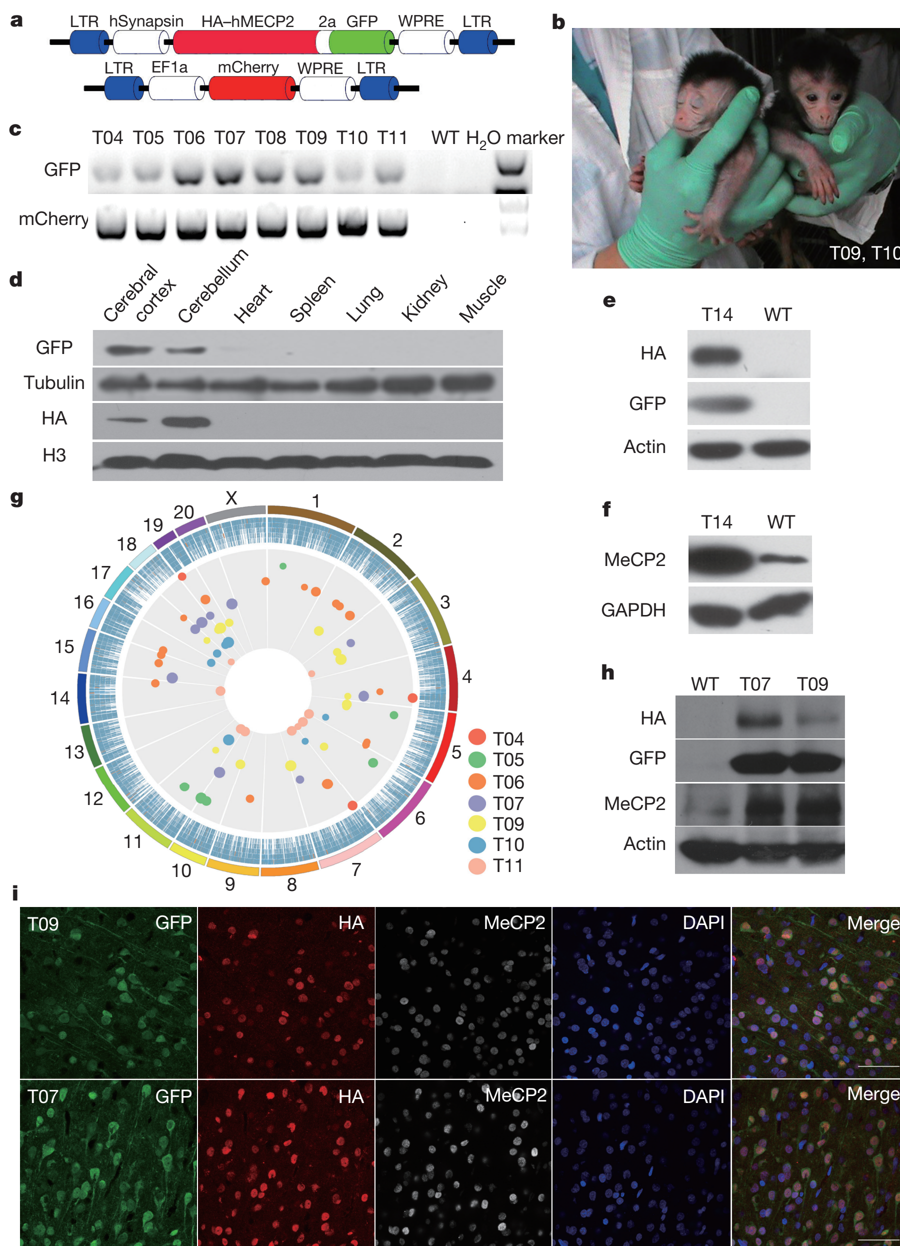


Figure 1 | Construction of *MECP2* transgenic monkey and brain-specific expression of transgenes. a, Top, lentiviral HA-hMECP2-2a-GFP cassette. Bottom, the EF1a-mCherry cassette. **b**, Image of newborn *MECP2* transgenic (TG) monkeys (T09 and T10). Photo credit: Y.W. **c**, PCR analysis showing the presence of transgenes (GFP, top; mCherry, bottom) in 8 live TG monkeys' genomes. **d**, Brain-specific transgene expression, shown by western blots of different tissues of T14. Top two panels: cytosolic fractions; bottom two panels: nuclear fractions, stained with antibodies indicated. Note the exposure times for transgene and the loading control were different (see Supplementary Fig. 1 for further comparison). H3, histone 3. **e**, **f**, **h**, Western blots showing expression of HA–MeCP2 and GFP in brain tissues of TG (T07, T09 and T14) and wild-type (WT) monkeys. For gel source data, see Supplementary Figs 1–3. **g**, Genome-wide distribution of transgenes in F_0 TG monkeys. Insertion sites (dots) distribute on various chromosomes (outermost circle). Sizes of dots are proportional to reads identified by deep-sequencing, colour-coded and aligned circularly for different monkeys (sample for T08 absent owing to preparation failure). **i**, Immunostaining of cortical sections of brains of T07 and T09 for GFP, HA, MeCP2 and DAPI. Scale bars, 50 μm.

lysates of T07 and T09 (Fig. 1h). Immunostaining of cortical slices of T07 and T09 showed that the MeCP2 and HA signals were co-localized (Fig. 1i), indicating expression of the *MECP2* transgene in the TG monkeys' brain.

Further transcriptome-wide analysis of the brain tissues was performed on four deceased TG (T14, T05, T07 and T09) and four WT monkeys using RNA-seq, based on the whole-genome sequencing data for cynomolgus monkeys[12–14]. We found 105 upregulated and 209 downregulated genes in TG monkeys (Extended Data Fig. 4d, e), with ≥2-fold change as compared to WT monkeys. Among them, 13 upregulated and 3 downregulated genes were also reported to exhibit similar changes in the *MECP2* transgenic mice[15].

Motor functions and responses to stress[16–18] were examined for 8 TG (T04–T11) and 8 WT monkeys (aged 12–18 months). First, we video-recorded the locomotion of each monkey alone for 20 min per day for 5 days, and found that four TG (T04, T05, T06 and T09) and two WT monkeys exhibited repetitive circular locomotion (in the same direction, at least three times; Supplementary Videos 1 and 2). The total time spent in circular locomotion during the observation period (average over 5 days) for all eight transgenic monkeys was significantly higher than that of eight WT monkeys (Fig. 2a, b). This difference

© 2016 Macmillan Publishers Limited. All rights reserved

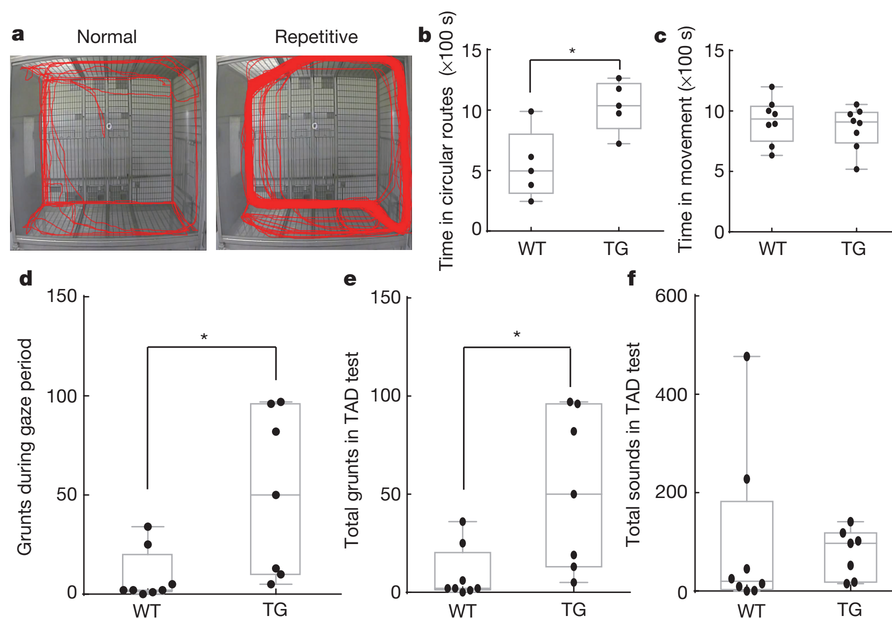


Figure 2 | Alterations in locomotion activity and increased anxiety in *MECP2* TG monkeys. a, Examples of movement trajectories (red traces) viewed from cage-top, showing normal activity (left) and repetitive circular routing (right). **b**, Boxplots of time spent in repetitive routing of TG and WT monkeys ($n = 8$ each), monitored for 20 min daily for 5 days. Each dot depicts data from 1 day (*$P = 0.014$, Student's t-test). **c**, Average total time spent in movements in 20-min period ($n = 8$ each). **d**, Results from threat-related anxiety and defence (TAD) test. Boxplots of numbers of grunts during the gaze period ($n = 8$, WT; $n = 7$, TG; *$P = 0.009$, Mann-Whitney U test) at 18 months of age. **e**, **f**, Total number of grunts (**e**) and all sounds (**f**) during the entire TAD tests (same monkey sets as in **d**; *$P = 0.014$, Mann–Whitney U test). Ends of whiskers represent the minimum and maximum of data points. The line within box represents the median (odd numbers of data points) or second quartile (even number of data points). The bottom and top edge of box represents the first and the third quartile, respectively.

was not due to hyperactivity of the TG monkeys, because the total time spent in locomotion was similar between the two groups during observation (Fig. 2c).

Anxiety-associated behaviours were found in autism patients and mouse models of MeCP2 overexpression[6,7]. We used the threat-related anxiety and defensive (TAD) behavioural test[9] to assay the vocalization responses of the monkeys to human gaze (Extended Data Fig. 5a). Typical sounds include grunt, coo and scream (Supplementary Audio 1–3), each was identified by its distinct signature in the sound spectrograph (Extended Data Fig. 6a–c). Notably, we found that at ~18 months of age, the total number of grunts made by the TG group during the gaze period and the entire TAD test was significantly higher than that of the WT group (Fig. 2d, e), with similar total numbers of sounds (grunt, coo and scream) per session produced by both groups (Fig. 2f and Supplementary Table 3). This increase in anxiety-related grunts was also found at 36 months of age (Extended Data Fig. 5b–d). Thus, *MECP2* TG monkeys showed increased levels of anxiety.

Impairment of social interaction is a hallmark of autism and *MECP2*-associated disorders. We examined the time monkeys sat together with apparent interactions, a prominent social behaviour in monkey colonies[19–21]. First, three groups of monkeys (at ~18 months of age, two TG and three WT in each group) were reared together for over 6 months (Supplementary Table 4a). We found that the average time a TG monkey sat together with another WT monkey within the group was significantly lower than that of the WT monkey (60-min daily observation for 5 days, Fig. 3a and Supplementary Video 3; no data for TG–TG interaction owing to the limited number of TG monkeys). The total time all TG monkeys sat with any other monkey (either TG or WT) was also slightly lower than for WT monkeys (Fig. 3a). Next (at ~24 months of age), we paired two female monkeys from different groups in a single cage (60 min daily for 5 days), and found that the interaction time of TG–TG pairs was significantly lower than that of WT–WT pairs (Fig. 3b, Supplementary Video 4 and Supplementary Table 4b). Pairing of unfamiliar male monkeys was not performed owing to their aggressive behaviours near adolescence. Finally (at 36 months of age), we paired familiar female and male monkeys from the same group, in which male pairs showed no aggressive interaction, and found that the male TG–TG pairs interacted less than that of the TG–WT pairs, whereas the difference between female TG–TG pairs and TG–WT pairs was not significant (Fig. 3c, d, Extended Data Fig. 7a–f and Supplementary Table 4c). All of these social interaction tests were performed with observers blinded to monkey genotypes. This apparent difference between male and female TG monkeys is reminiscent of the finding that *MECP2* duplication syndrome show more severe autism-related symptoms in male patients[6].

Cognitive function tests were performed using the Wisconsin general test apparatus (WGTA)[22–26]. During adaptation, discrimination and reversal steps of black/white tests, both the WT and TG groups passed each step with a similar average time course, but the TG monkeys exhibited much larger variability and one (T11) failed to pass the black/white reversal step (and was thus dropped from subsequent tests) (Extended Data Fig. 8a, b and Supplementary Table 5). In the Hamilton searching tests (adaptation, searching, set-breaking and forced set-breaking), the TG group showed a slightly slower learning in the forced set-breaking step (Extended Data Fig. 8c, d). Finally, the two groups showed no significant difference in the average performance in the reward-shape association learning test (Extended Data Fig. 9a–c and Supplementary Video 5). However, three out of seven TG monkeys showed marked left-side preference, regardless of the left or right location of the reward, a behaviour that was not observed in the WT monkeys (Fig. 3e, f and Supplementary Video 6). Thus, *MECP2* transgene expression resulted in some abnormalities in cognitive behaviours.

We further examined the germline transmission of *MECP2* TG monkeys, in view of previous lentiviral-based transgenic monkey experiments[27,28]. To facilitate the reproduction of TG monkeys, we used a recently developed testicular tissue xenografting method (see Methods). In brief, one testicle was obtained from T07 at 27 months of age, and pieces of the testicle tissue were xenografted subcutaneously in nude mice. Mature motile sperm were obtained from the xenografts after 10 months and used for ICSI on 176 monkey oocytes.

© 2016 Macmillan Publishers Limited. All rights reserved

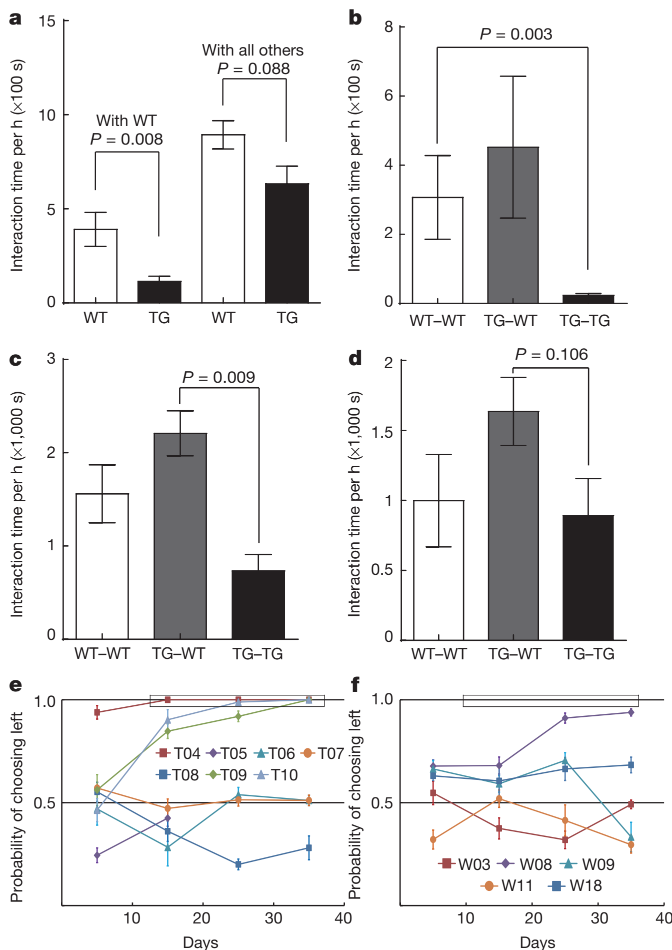


Figure 3 | Impaired social interaction and cognitive functions in *MECP2* TG monkeys. a, Interaction at 18 months of age among monkeys within the same group ('familiar', reared together for 6 months), defined by the average time a WT or a TG monkey sat together (for >3 s) with a WT monkey from the same group (left), or with any other monkey (right) during a 60-min observation period ($n = 6$, TG; $n = 9$, WT). **b**, Interaction time between a pair of female monkeys from a different group ('unfamiliar', reared separately for 12 months) at 24 months of age, for WT–WT, WT–TG and TG–TG pairs ($n = 4$, TG; $n = 12$, WT). **c**, **d**, Interaction time between familiar monkey pairs at 36 months of age (**c**, male pairs, $n = 3$, TG; $n = 8$, WT; **d**, female pairs, $n = 4$, TG; $n = 11$, WT) from the same group reared together for 6 months. All P values are from Mann–Whitney U test. **e**, **f**, Data from WGTA test ($n = 7$, TG; $n = 5$, WT), showing distinct left bias in the probability of left or right choice in three TG monkeys. Boxed areas denote >95% probability of choosing left side. Error bars denote s.e.m.

Implantation of 95 fertilized zygotes into 22 surrogates resulted in five F_1 offspring (TF1-1 to TF1-5; 4 live births; Fig. 4a and Extended Data Table 1b). The PCR analysis showed that all five F_1 monkeys carried the human *MECP2* transgene (Fig. 4b), and western blotting showed the expression of HA–MeCP2 and GFP in brain lysates of TF1-1 (deceased 3 days after birth) (Fig. 4c). Thus, the transgenes were expressed in the F_1 offspring. Deep-sequencing further showed the genomic integration sites of the HA-hMECP2-2a-GFP and mCherry transgenes in F_1 TG monkeys (Fig. 4d and Supplementary Table 6). As expected, transgenes in F_1 TG monkeys (TF1-1 to TF1-5) were mostly distributed in chromosomes that were a subset of transgene-containing chromosomes of the paternal monkey T07 (Fig. 4d), showing Mendelian segregation of transgenes among the F_1 progeny. Moreover, long-terminal repeat (LTR) insertion sites were consistent with AccuCopy-identified copy numbers, confirming the germline transmission of transgenes (Fig. 4e).

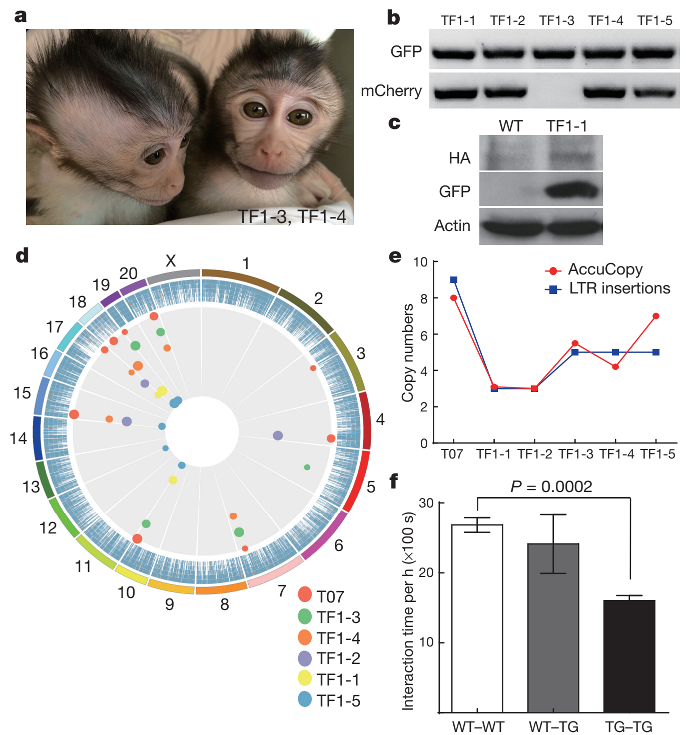


Figure 4 | Generation of F_1 progeny of *MECP2* TG monkey. a, Image of two newborn F_1 TG monkeys (TF1-3 and TF1-4). Photo credit: Y.-H.N. **b**, PCR analysis showing the presence of transgenes (GFP and mCherry) in five F_1 monkeys' genomes. **c**, Western blots showing expression of the HA–MeCP2 and GFP in F_1 monkeys' brain. For gel source data, see Supplementary Fig. 4. **d**, Genomic-wide distribution of transgenes in T07 and five F_1 offspring. Insertion sites (dots) distribute on various chromosomes (outermost circle). Sizes of dots are proportional to reads identified by deep-sequencing, colour-coded and aligned circularly for different monkeys. **e**, Copy numbers obtained with AccuCopy (red, *MECP2* and mCherry) are consistent with LTR insertion numbers identified by deep-sequencing (blue). **f**, Social interaction time between *MECP2* TG F_1 and WT monkeys from a different group ('unfamiliar pairing') at 11 months of age, for WT–WT, WT–TG and TG–TG pairs (P value, Student's t-test) ($n = 4$, TG; $n = 6$, WT). Error bars denote s.e.m.

Finally, we examine whether F_1 TG monkeys may also show defects in social interaction. We set up two groups of monkeys at 11 months of age (two TG and three WT in each group; Supplementary Table 7), then paired two monkeys from two different groups in a single cage (60 min daily for 5 days), observed in a blinded manner. We found that the TG–TG pairs showed a significantly lower total interaction time than the WT–WT pairs (Fig. 4f), although in general young monkeys exhibited more frequent interaction than older monkeys. Thus, defective social behaviours were inherited in the F_1 generation of *MECP2* TG monkeys.

In summary, we have generated transgenic cynomolgus monkeys by using lentiviral infection of monkey oocytes. These TG monkeys showed an increased frequency of repetitive circular locomotion, increase anxiety, reduced social interaction and relatively weak cognitive phenotypes. Overall, we found no evidence of correlation between the copy number of transgenes and the extent of behavioural abnormalities, presumably owing to the low sample number with each copy number, and the possibility of nonspecific effects of gene transfer on behaviours could not be excluded. Importantly, we generated five F_1 TG offspring from one founder TG monkey, confirming the feasibility of germline transmission of lentiviral-based genetic engineering in monkeys. Together with recent progress in applying new gene-editing methods in monkey[29,30], our findings pave the way for the efficient use of genetically engineered macaque monkeys for studying brain disorders.

© 2016 Macmillan Publishers Limited. All rights reserved

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 31 May; accepted 14 December 2015.
Published online 25 January 2016.

1. Meehan, R. R., Lewis, J. D., McKay, S., Kleiner, E. L. & Bird, A. P. Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell* **58,** 499–507 (1989).
2. Nan, X., Campoy, F. J. & Bird, A. MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell* **88,** 471–481 (1997).
3. Young, J. I. *et al.* Regulation of RNA splicing by the methylation-dependent transcriptional repressor methyl-CpG binding protein 2. *Proc. Natl Acad. Sci. USA* **102,** 17551–17558 (2005).
4. Cheng, T. L. *et al.* MeCP2 suppresses nuclear microRNA processing and dendritic growth by regulating the DGCR8/Drosha complex. *Dev. Cell* **28,** 547–560 (2014).
5. Amir, R. E. *et al.* Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nature Genet.* **23,** 185–188 (1999).
6. Ramocki, M. B. *et al.* Autism and other neuropsychiatric symptoms are prevalent in individuals with MeCP2 duplication syndrome. *Ann. Neurol.* **66,** 771–782 (2009).
7. Samaco, R. C. *et al.* Crh and Oprm1 mediate anxiety-related behavior and social approach in a mouse model of MECP2 duplication syndrome. *Nature Genet.* **44,** 206–211 (2012).
8. Collins, A. L. *et al.* Mild overexpression of MeCP2 causes a progressive neurological disorder in mice. *Hum. Mol. Genet.* **13,** 2679–2689 (2004).
9. Kalin, N. H. & Shelton, S. E. Defensive behaviors in infant rhesus monkeys: environmental cues and neurochemical regulation. *Science* **243,** 1718–1721 (1989).
10. Nakagawa, T. *et al.* Generation of lentiviral transgenic rats expressing glutamate receptor interacting protein 1 (GRIP1) in brain, spinal cord and testis. *J. Neurosci. Methods* **152,** 1–9 (2006).
11. Frye, R. E., Melnyk, S. & Macfabe, D. F. Unique acyl-carnitine profiles are potential biomarkers for acquired mitochondrial disease in autism spectrum disorder. *Transl. Psychiatry* **3,** e220 (2013).
12. Yan, G. *et al.* Genome sequencing and comparison of two nonhuman primate animal models, the cynomolgus and Chinese rhesus macaques. *Nature Biotechnol.* **29,** 1019–1023 (2011).
13. Higashino, A. *et al.* Whole-genome sequencing and analysis of the Malaysian cynomolgus macaque (*Macaca fascicularis*) genome. *Genome Biol.* **13,** R58 (2012).
14. Huh, J. W. *et al.* Large-scale transcriptome sequencing and gene analyses in the crab-eating macaque (*Macaca fascicularis*) for biomedical research. *BMC Genomics* **13,** 163 (2012).
15. Chahrour, M. *et al.* MeCP2, a key contributor to neurological disease, activates and represses transcription. *Science* **320,** 1224–1229 (2008).
16. Ramiro, L. S., Madrid, B. J. & Brown, D. W. Adverse childhood experiences (ACE) and health-risk behaviors among adults in a developing country setting. *Child Abuse Negl.* **34,** 842–855 (2010).
17. Drago, L. & Thierry, B. Effects of six-day maternal separation on tonkean macaque infants. *Primates* **41,** 137–145 (2000).
18. Suomi, S. J. Early determinants of behaviour: Evidence from primate studies. *Br. Med. Bull.* **53,** 170–184 (1997).
19. Feng, X. L. *et al.* Maternal separation produces lasting changes in cortisol and behavior in rhesus monkeys. *Proc. Natl Acad. Sci. USA* **108,** 14312–14317 (2012).
20. Bauman, M. D., Lavenex, P., Mason, W. A., Capitanio, J. P. & Amaral, D. G. The development of social behavior following neonatal amygdala lesions in rhesus monkeys. *J. Cogn. Neurosci.* **16,** 1388–1411 (2004).
21. Emery, N. J. *et al.* The effects of bilateral lesions of the amygdala on dyadic social interactions in rhesus monkeys (*Macaca mulatta*). *Behav. Neurosci.* **115,** 515–544 (2001).
22. Golub, M. S., Hogrefe, C. E. & Germann, S. L. Iron deprivation during fetal development changes the behavior of juvenile rhesus monkeys. *J. Nutr.* **137,** 979–984 (2007).
23. Makori, N., Watson, R. E., Hogrefe, C. E., Lalayeva, N. & Oneda, S. Object discrimination and reversal learning in infant and juvenile non-human primates in a non-clinical laboratory. *J. Med. Primatol.* **42,** 147–157 (2013).
24. Sackett, G., Ruppenthal, G., Hewitson, L., Simerly, C. & Schatten, G. neonatal behavior and infant cognitive development in rhesus macaques produced by assisted reproductive technologies. *Dev. Psychobiol.* **48,** 243–265 (2006).
25. Harlow, H. F. The development of learning in the rhesus monkey. *Am. Sci.* **47,** 458–479 (1959).
26. Ha, J. C., Mandell, D. J. & Gray, J. Two-item discrimination and Hamilton search learning in infant pigtailed macaque monkeys. *Behav. Processes* **86,** 1–6 (2011).
27. Sasaki, E. *et al.* Generation of transgenic non-human primates with germline transmission. *Nature* **459,** 523–527 (2009).
28. Moran, S. *et al.* Germline transmission in transgenic Huntington's disease monkeys. *Theriogenology* **84,** 277–285 (2015).
29. Liu, H. *et al.* TALEN-mediated gene mutagenesis in rhesus and cynomolgus monkeys. *Cell Stem Cell* **14,** 323–328 (2014).
30. Liu, Z. *et al.* Generation of a monkey with *MECP2* mutations by TALEN-based gene targeting. *Neurosci. Bull.* **30,** 381–386 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank M.-m. Poo for comments on the manuscript, Y.-Z. Li, Y.-Z. Lu, F. Liu and X. Zhou for maintaining monkey colony, W. Lu, X.-H. Guo and Y. F. Zhou of Fudan Children's Hospital for assistance in mass spectrometry and electroencephalogram, D. Chen and B. Zhang of Novel Bioinformatics for RNA-seq data analysis, and C.-H. Li of Shanghai Geneskies Company for analysis of genomic integration sites. This work was supported by CAS Strategic Priority Research Program (XDB02050400), the MoST 973 Program (2011CBA00400), NSFC grants (91432111, 91232712 and 81527901), National Key Technology R&D Program of China 2014BAI03B00, Shanghai City Committee of Science and Technology Project 14140900100.

Author Contributions Z.Q. and Q.S. conceived and supervised the project. T.-L.C. constructed the lentiviral constructs. Q.S. and Z.L. performed the cynomolgus oocytes preparation and injection. Y.-J.C., Y.W., C.-C.Z., Y.-H.N. and Z.L. contributed to monkey reproductive experiments. Y.-F.Z. performed PCR-based genotyping experiments. Z.-F.C., W.-J.B., X.-D.Z. and X.Y. performed immunohistochemistry and AccuCopy experiments. C.C., B.L., X.S. and Z.-Q.X. performed western blot experiments. X.L. and J.-J.W. performed behavioural analysis. J.-T.Z. and N.G. performed WGTA tests. W.-H.Z. and X.X. contributed to metabolic measurements and behavioural analysis. T.-L.C. and X.L. performed genomic integration sites analysis based on deep-sequencing. J.X., L.Z. and F.Z. helped with identification of genomic integration sites of transgenes. Z.Q. wrote the manuscript.

Author Information The raw sequence and processed data have been submitted to the NCBI Gene Expression Ominbus (GEO) under accession number GSE57974. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to Z.Q. (zqiu@ion.ac.cn) or Q.S. (qsun@ion.ac.cn).

© 2016 Macmillan Publishers Limited. All rights reserved

METHODS

Animal ethics statement. The use and care of animals complied with the guideline of the Biomedical Research Ethics Committee at the Shanghai Institutes for Biological Science (CAS), which approved the application entitled 'Reproductive physiology of cynomolgus monkey and establishment transgenic monkey' (#ER-SIBS-221106P).

Collection of oocytes, gene delivery and embryo construction. Laparoscopy was used for oocyte collection. Oocytes were aspirated from follicles 2–8 mm in diameter, about 32–36 h after hCG stimulation[31]. The collected oocytes were cultured in the pre-equilibrated maturation medium[32]. Metaphase II arrested oocytes were selected for perivitelline space injection[32] of lentiviruses and ICSI. The lentivirus concentration for injection was 1×10^{10} viral genome (vg) per ml. After microinjection, the oocytes were cultured in the maturation medium at 37 °C (in 5% CO_2) for about 1 h, until fertilization by ICSI. Monkey semen was collected by penile electro-ejaculation. For ICSI, a single sperm was immobilized and aspirated with the tail first. A single oocyte was fixed by the holding pipette, and the injection pipette was pushed through the zona pellucida and subsequently through the oolemma to release the spermatozoon[32]. After ICSI, the oocytes were cultured in pre-equilibrated Hamster Embryo Culture Medium 9 (HECM-9) at 37 °C (in 5% CO_2) until the next morning[33,34].

Selection of surrogate females and embryo transfer. Menstrual cycles of females were recorded daily. To synchronize the developmental stage of embryos with the recipient, monkeys were chosen for tubal embryo transfer at 0–3 days after ovulation, and a stigma or a new corpus luteum on the ovary could be observed by laparoscopy. About 2–3 pronuclear-stage embryos were selected for tubal transfer to each surrogate female[31].

Genotyping PCR. Hair-root samples collected from newborn monkey pups were used to extract DNA. Samples were digested by proteinase K overnight at 65 °C and precipitated for DNA and PCR with specific primers again GFP and mCherry were used for initial genotyping analysis as follows: mCherry-R: 5′-TGCTTGATCTCGCCCTTCAG-3′, mCherry-F: 5′-GCCATCATCAAGGAGTTCATGC-3′; GFP-F: 5′-AAGTTCATCTGCACCACCG-3′, GFP-R: 5′-TCCTTGAAGAAGATGGTGCG-3′.

Southern blot analysis. A total of 15 μg of genomic DNA was prepared and digested with BamHI and EcoRI, which released transgenes. Genomic DNAs were separated with 1% agarose gel and transferred to Nippon N+ membrane (GE). DNA probes from hMECP2-2a-GFP was prepared using ready-to-go DNA label kit (279240D-20, GE Life Sciences). ^{32}P-labelled probes were hybridized with blots of genomic DNAs and exposed to phosphor-imager after extensively washing.

Euthanasia procedure for sick and aborted newborn monkeys. Decisions of whether euthanasia procedures would be carried out for sick or aborted newborn monkeys are made by veterinarians, after consulting with principal investigators and followed the approved protocol (#ER-SIBS-221106P). Aborted or sick *MECP2* TG and WT monkeys were deeply anaesthetized with ketamine hydrochloride (5–10 mg kg^{-1}) to avoid possible pain and then perfused with 0.9% saline with 2–4% paraformaldehyde (PFA) for further immunohistochemistry experiments. The procedure is approved by the Biomedical Research Ethics Committee at the Shanghai Institutes for Biological Science (CAS), described in the protocol entitled 'Reproductive physiology of cynomolgus monkey and establishment transgenic monkey' (#ER-SIBS-221106P).

Immunohistochemistry of transgenic monkey tissues. After perfusion, the hemispheres of the brain were dissected, cut in to small blocks, fixed with 4% PFA in phosphate buffer, and equilibrated in 30% sucrose. Fixed and equilibrated brain tissue blocks were cut into 30-μm cortical sections with a Microm HM525 cryostat. Sections were washed for 5 min in PBS containing 5% bovine serum albumin (BSA) and 0.3% Triton X-100, and incubated with primary antibodies (in PBS with 3% BSA and 0.3% Triton X-100) overnight at 4 °C and subsequently with corresponding secondary antibodies (Alexa-Fluor-conjugated, Invitrogen, at 1:1,000). DAPI was used to label the nuclei and sections were mounted with 75% glycerol. Other antibodies used: HA antibody (Covance, MMS-101R), NeuN antibody (Millipore, MAB377), MeCP2 antibody (Cell Signaling, 3456S) and GFP antibody (Abcam, ab6673).

Measurement of *MECP2* transgene copy numbers using AccuCopy assay. Four sets of primers targeted to *MECP2* were designed. One set (mecp2_1) was a cross-intron primer targeted to transgenic cDNA fragments representing the copy number of transgenic DNA; the second (mecp2_2) was targeted to one exon of transgenic cDNA fragments representing the total *MECP2* copy number; and the other two primer sets (mecp2IN_1 and mecp2IN_2) were targeted to introns of monkey *MECP2* gene representing the endogenous *MECP2* copy number. Two sets of *EGFP* primers (EGFP_1 and EGFP_2) were designed to verify the copy number of the transgene, and one set of mCherry primers was designed as negative control. The copy number of these DNA fragments was measured using custom-designed Multiplex AccuCopyTM Kit (Geneskies Biotechnologies, CN0105). The copy number of these target DNA fragments was measured using custom-designed Multiplex AccuCopy kit (Geneskies Biotechnologies, CN0105). For each DNA fragment amplified, a piece of synthesized competitive double-stranded DNA of known concentration and with insertions or deletions of a few base pairs was added to the PCR reaction mix. Each PCR reaction was carried out by mixing the synthesized competitive double-stranded DNAs for target and reference genes (*POP1*, *RPP14* and *POLR2A*) together with a defined amount of sample DNAs. A multiplex competitive PCR was then performed to simultaneously amplify all reference and target genes from both sample and competitive DNAs using multiple fluorescence-labelled primer pairs. In brief, the 20-μl PCR reaction for each sample contained 1× AccuCopy PCR Master Mix, 1× Fluorescence Primer Mix, 1× Competitive DNA mix and ~10 ng sample DNA. The PCR program used was: 95 °C for 10 min; 11 cycles of 94 °C for 20 s, 65 °C–0.5 °C/cycle 40 s, 72 °C for 1.5 min; 24 cycles of 94 °C for 20 s, 59 °C for 30 s, 72 °C for 1.5 min; 60 °C for 60 min. PCR products were diluted 20-fold before loaded on ABI3730XL sequencer (Applied Biosystems) to separate amplicons of different sizes by capillary electrophoresis. Raw data were analysed using GeneMapper4.0, and the peak ratios of sample DNA to competitive DNA (S/C ratio) for all target and reference fragments were exported to Excel. The S/C ratio of each target fragment was first normalized to the S/C ratio of the reference genes, and then further normalized to the median copy number of the entire data set. The final normalized ratio was averaged for each *MECP2* primer and *EGFP* primer, and the similarity between the two ratio further confirmed the copy number of the transgene.

Primer sequences: mecp2_1-for 5′-CGCTCTGCTGGGAAGTATGATG-3′, mecp2_1-rev 5′-GGGATGTGTCGCCTACCTTTTC-3′; mecp2_2-for 5′-AAGCCCAAATCTCCCAAAGCTC-3′, mecp2_2-rev 5′-TTCCCAGGACTTTTCTCCAGGAC-3′; mecp2IN_1-for 5′-GCAAGGTTTGGCTGAAGGAGAA-3′, mecp2IN_1-rev 5′-GAGCACACCCCACAGCAGTAAA-3′; mecp2IN_2-for 5′-TTGGGCTTGAACTCCAGACCTC-3′, mecp2IN_2-rev 5′-GGTGGTGCCACTGACTGAGAAG-3′; EGFP_1-for 5′-TACGGCAAGCTGACCCTGAAGT-3′, EGFP_1-rev 5′-CTGCTTCATGTGGTCGGGGTAG-3′; EGFP_2-for 5′-CTGCCCGACAACCACTACCTG-3′, EGFP_2-rev 5′-GAACTCCAGCAGGACCATGTGA-3′; POP1-for (internal control) 5′-AGAAATGCTTGGGCCTGTTACG-3′; POP1-rev (internal control) 5′-CACAGCTGCCTGCTCTCAGAAG-3′; RPP14-for (internal control) 5′-TTTTTTGGTGCCTGCATGTTTG-3′, RPP14-rev (internal control) 5′-AGCGATGGCTGGAAATTAGTGG-3′; POLR2A-for 5′-GGTGAAGTACGACGCGACTGTG-3′, POLR2A-rev 5′-AGGTTTCAGCGTAGCCAGGTTC-3′.

Preparation of lentivirus. Lentiviruses were produced by standard protocols and provided at a titre of 10^{10} vg ml^{-1} by the Shanghai SBO Medical Biotechnology Co. Ltd.

Identification of genomic integration sites of transgenes. A total of 2 μg genomic DNA was used to construct a DNA library for each case[35–38]. Sequencing linkers were further added onto genomic segments (length around 500–700 base pairs (bp)) (Extended Data Fig. 1b). After end repairing and 3′ A-adding, the fragmented DNAs were ligated with Y-shape adaptor. Amplification was performed with the adaptor primers. Asymmetry-primer PCR (APP) was used to enrich the viral integration sites in each library. The APP method includes two PCR systems. The first PCR system includes only LTR specific primer. After 12 cycles of linear amplification, adaptor specific primer was added in the PCR system followed by 12 cycles of exponential amplification. PCR products were purified using 0.7 × AMPure beads (Beckman, A63882). The second PCR system uses a pair of primers nest the primers in the first PCR system. After 12 cycles of linear amplification and 15 cycles of exponential amplification, the PCR products of 500–700 bp in size were isolated by agarose gel electrophoresis before being used to construct libraries with Illumina paired-end adapters according to the manufacturer protocol and sequenced by Illumina MiSeq V3 (2 × 300 base paired ends). Only the paired-end reads showing the fusions of viral sequences and the cynomolgus (*Macaca fascicularis*) genome segments were selected, in which two mismatches were allowed. The reads showing the same integration position were merged and treated as a unique integration site. Experiments were repeated three times independently with different sequencing linkers. Determination of insertion sites is under the following criteria: (1) total insert numbers are greater than 100 times after three experiments; (2) being detected at least twice after three experiments. Cynomolgus monkey genome is used in the following database: http://www.ncbi.nlm.nih.gov/genome/?term=crab+eating+monkey. Target sequences containing LTR of transgene cassettes and genomic segments flanking the transgenes were analysed (Supplementary Tables 2 and 6).

Primer sequences: LTR_1: 5′-CTTGCCTTGAGTGCTTCAAGTAGT-3′; LTR_2: 5′-TGCCCGTCTGTTGTGTGACTCT-3′; YP_1: 5′-GGATAGCGACGCACGGAACTCT-3′; YP_2: 5′-CTCCATCTCATCCTGCGTGTC-3′.

© 2016 Macmillan Publishers Limited. All rights reserved

Western blotting. Monkey brain tissues were homogenized in RIPA buffer (containing 50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1% Triton X-100, 0.1% SDS, 1% sodium deoxycholate, protease inhibitor cocktail and phosphatase inhibitor cocktail) on ice and then centrifuged at 1,000*g* for 10 min at 4 °C. The supernatant was stored at −80 °C until use. Protein concentration was measured with BCA method. Approximately 30 μg protein of each sample was loaded in on 10% SDS–PAGE and run at 120 V constant voltage. A constant current of 0.36 mA was used for transblotting. Blots were probed with primary antibodies (1:1,000) overnight at 4 °C. After washing three times, blots were then incubated with goat anti-rabbit secondary antibody (1:3,000) at room temperature for 2 h. Chemiluminescence was used to visualize protein bands. Antibodies used: HA antibody (Abcam, ab9110), MeCP2 antibody (Cell Signaling, 3456S) and GFP antibody (Invitrogen, A11122).

Mass spectrometry analysis. Fresh whole blood from upper arms of monkeys was taken by a professional veterinarian in the morning before feeding. Whole blood (200–400 μl) was dropped onto filter paper immediately. After air drying, filter papers were store at −20 °C before mass spectrometry analysis. API2000 from AB SCIEX were used for analysing fatty acid and amino acids. Data were obtained from three rounds of blood collections independently.

Behaviour observation and analysis. Behaviour observation and analysis were performed by two independent trained observers, with demonstrated inter-observer reliability of at least 80%. All observers were blinded to the genotypes of the monkeys. The dimension of cages using for living and behaviour monitoring is $1.5 \times 1 \times 1.1$ m.

Analysis of locomotive behaviour. Monkeys individually, were observed alone in an observation cage ($1.5 \times 1 \times 1.1$ m) after they had been accustomed to community living following weaning. The observation cage was similar to their home cage. All locomotion behaviours were video-record without interruption for 20 min each day for 5 days. Data from 5 days were pooled.

Analysis of social interaction behaviours. Social behaviours of TG and WT monkeys with familiar and unfamiliar monkeys were studied by examining the interactions of monkey from the same and different home cage, respectively. To study the interaction with familiar monkeys, we housed three groups of monkeys, each consisting three WT and two TG monkey of the same age, in three separate cages for 6 months before the observation (at about 1.5 years old). In this analysis, the observer followed the time each monkey spent sitting together with another monkey for a duration of 1 h each day for 5 consecutive days. We defined that two monkeys sat together by obvious interactions between the two for more than 3 s, during which the monkeys may exhibit touching and grooming behaviours or lean against each other. To study interaction with unfamiliar monkeys, we regrouped the females from same cohorts after the above observation for another 8 months in four separate cages (see Supplementary Table 4a, b). (Males were kept together separately owing to their proximity to sexual maturity, thus not used for observation). For each observation of social interaction, we paired two monkeys from different group and observation was made in the same manner as that described above for the interaction between familiar monkeys.

Analysis of social interaction behaviours of F_1 generations. To study the interaction with F_1 TG monkeys, we housed two groups of monkeys (group info see Supplementary Table 7), each consisting of three WT and two TG F_1 monkeys of similar age (at 10–11 months old), in two separate cages before the observation. In this analysis, the observer followed the time each monkey spent sitting together with another monkey for a duration of 1 h each day for 5 consecutive days.

Analysis of TAD behaviours. The TAD behavioural model was used to assay the monkey's response to human gaze (Extended Data Fig. 5a). In each session of observation, an individual monkey from either the transgenic or WT group was placed in an observation cage ($1.5 \times 1 \times 1.1$ m), and allowed to adapt to the cage alone for 9 min. An observer then sat in front of the cage at a distance of 2 m, showing the face profile to the monkey without eye contact for 9 min ('non-gaze period'). This was followed by the relaxation period (3 min) without the human presence, and the 'gaze period' (9 min) in which the observer sat in front of the cage and gazed at monkey with a neutral face. Behaviour and vocalizations were recorded on videotape[39–41].

WGTA. WGTA tests were performed on 8 TG and 6 WT monkeys at the age of 1.5 years, in accordance to WGTA protocol[25,26], by trained technicians. The WGTA apparatus includes a testing box that for observing subject's activity, a presentation board with food wells for reward placing, a trial door and an access door connected by pulley cord to separate the subject and presentation board, and a camera for recording. All tests were carried out in a quiet and standard lighted room.

Black/white test. This test includes three stages: adaptation, discrimination and reversal. For the adaptation step, each monkey was tested for the ability to take the food reward on the presentation board that was placed by experimenter. Before the adaptation step, the monkey needs to pass several pre-test steps: the reward was placed in front of the food well, in the food well, in the food well next to the adaptation block, and in the food well with half covered by the adaptation block. Finally for adaptation step, the monkey had to take the food in the food well with the block covered completely. Each monkey received a maximum of 25 trials per day, and was considered to be passed when showing correct responses on 23 out of 25 trials. During the discrimination step, each monkey needed to choose the only reward in the food well that was covered by either a black or white block with an empty well covered the opposite colour. The same monkey was always rewarded with either black or white but with random location, with assignment of monkeys by the Gellerman order. Each monkey received 25 trials per day and was considered to be passed when showing correct responses on 23 out of 25 trials. For the reversal step, the procedure was the same as discrimination step, except that the monkey was rewarded black if white was rewarded during the discrimination step, and vice versa.

Hamilton search test. This test includes four steps: adaptation, Hamilton search, Hamilton search set-breaking, and Hamilton search forced set-breaking. The adaptation step was similar to that for black/white test, with the same criterion for passing. For the Hamilton search step, four little boxes that represented the different positions from experimenter's left to right on the presentation board were used for testing. The only reward was randomly placed in one of the four closed boxes in each trial. Monkey was allowed to find the reward from these four closed boxes. One trial was terminated when the monkey open the correct box. Each monkey performed 25 trials per day for 5 consecutive days. For the Hamilton search set-breaking step, the box that was the least preferred was first determined from the above step, and was always rewarded when chosen by the monkey. One trial was terminated when the subject open the correct box. Each monkey performed 25 trials per day for 5 days. For the Hamilton search forced set-breaking step, the procedure was the same as the set-breaking test, except that the monkey was allowed to make only one choice for finding the reward that placed in the least preferred box. The monkey was scored for the rate of correct choice over 25 trials each day for 5 consecutive days.

Learning set test. The monkeys were tested for the ability to distinguish 240 pairs of toys. The toys in each pair were labelled A or B to cover the two food wells, one of which had food. For each monkey, either A or B was always rewarded. Each pair of toys was presented for 6 trials and 6 pairs were tested each day. Six different pairs were used for different days, with the test lasting 8 weeks until all 240 pairs were used. The monkey was scored for the rate of correct choice, averaged over 180 trials (5 days).

Transcriptome analysis for transgenic monkey. Total RNA was extracted from three independent pieces of cortical tissues from brains of T05, T07, T09 and T14 and four WT monkeys by Trizol reagent (Invitrogen) separately. The RNA quality was checked by Bioanalyzer 2200 (Aligent) and kept at −80 °C. The RNA with RIN (RNA integrity number) > 8.0 is acceptable for cDNA library construction. RNA-seq and bioinformatic data analysis were performed by Shanghai Novelbio Ltd.

Library construction and RNA-seq. The cDNA libraries for single-end sequencing were prepared using Ion Total RNA-Seq Kit v2.0 (Life Technologies) according to the manufacturer's instructions. The cDNA libraries were then processed for the proton sequencing process according to the commercially available protocols. Samples were diluted and mixed, the mixture was processed on a OneTouch 2 instrument (Life Technologies) and enriched on a OneTouch 2 ES station (Life Technologies) for preparing the template-positive Ion PI Ion Sphere Particles (Life Technologies) according to Ion PI Template OT2 200 Kit v2.0 (Life Technologies). After enrichment, the mixed template-positive Ion PI Ion Sphere Particles of samples was loaded on to 1 P1v2 Proton Chip (Life Technologies) and sequenced on Proton Sequencers according to Ion PI Sequencing 200 Kit v2.0 (Life Technologies).

RNA-seq mapping. Before read mapping, clean reads were obtained from the raw reads by removing the adaptor sequences, reads with >5% ambiguous bases (noted as *N*) and low-quality reads containing more than 20% of bases with qualities of <13. The clean reads were then aligned to crab eating macaque genome (version: Mfa5.0) using the MapSplice program (v2.1.6). In alignment, preliminary experiments were performed to optimize the alignment parameters (-s 22 -p 15–ins 6–del 6–non-canonical) to provide the largest information on the AS events[42].

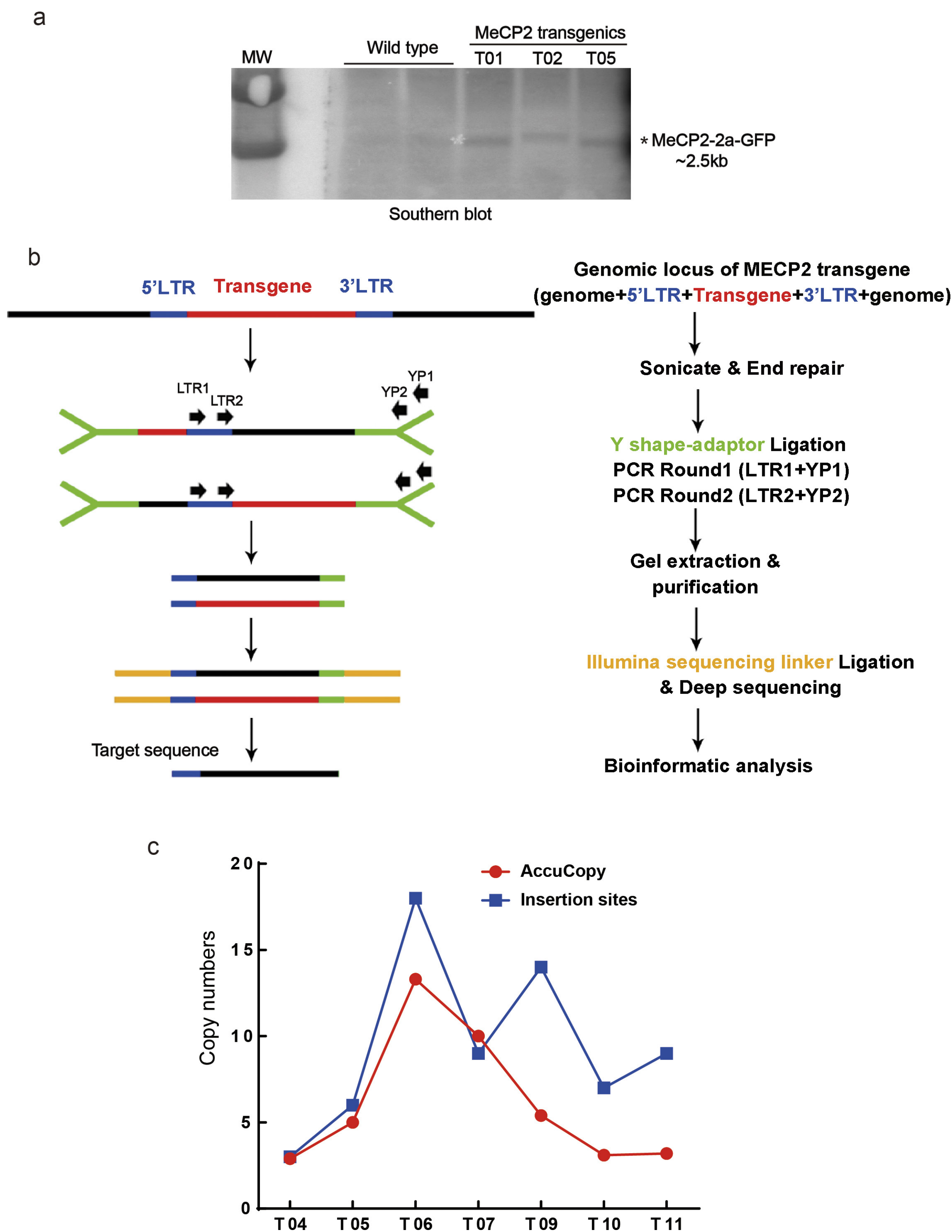
Dif-Gene-Find er-t. We applied DEseq algorithm to filter the differentially expressed genes, after the significant analysis and false discovery rate (FDR) analysis under the following criteria: (1) fold change > 1.5 or < 0.667; (2) FDR < 0.05 (ref. 43). A Volcano plot was drawn by *P* value based on the differential gene analysis, and the colour was determined by the filtering criteria (red, $\log_{10}(P \text{ value}) > 1.5$; blue, $\log_{10}(P \text{ value}) < 1.5$; black, $\log_2(\text{FC(TG/WT)}) < \pm 0.5$).

Generation of F_1 offspring of transgenic monkeys. The F_1 offspring was generated by ICSI using sperms obtained from testicular tissue xenografts of the T07 monkey. The method of testicular xenografting greatly shortened the time required for sexual maturation of TG monkey[44].

© 2016 Macmillan Publishers Limited. All rights reserved

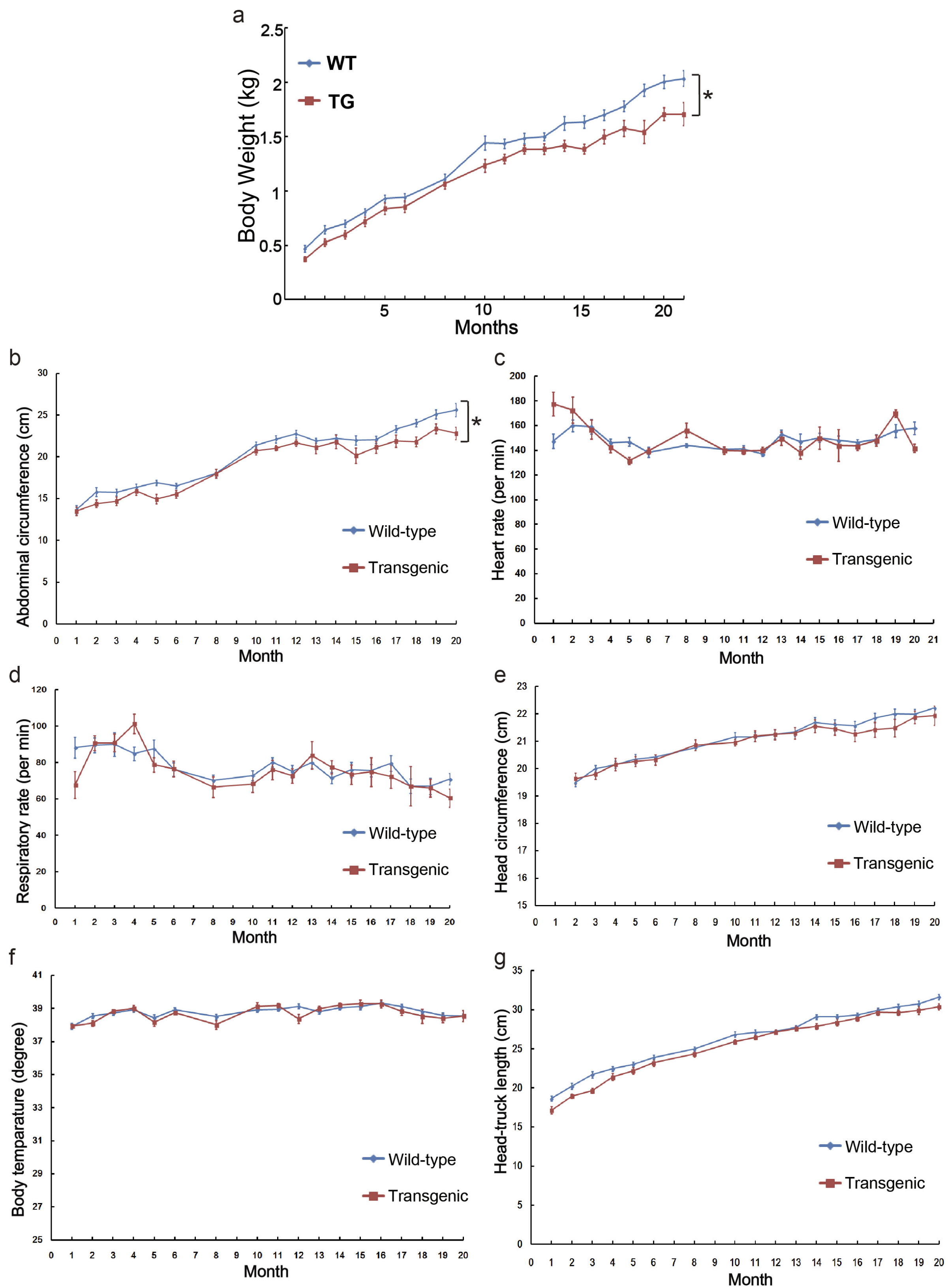
31. Sun, Q. *et al.* Efficient reproduction of cynomolgus monkey using pronuclear embryo transfer technique. *Proc. Natl Acad. Sci. USA* **105,** 12956–12960 (2008).
32. Chan, A. W. & Yang, S. H. Generation of transgenic monkeys with human inherited genetic disease. *Methods* **49,** 78–84 (2009).
33. Zheng, P., Bavister, B. D. & Ji, W. Z. Energy substrate requirement for in vitro maturation of oocytes from unstimulated adult rhesus monkeys. *Mol. Reprod. Dev.* **58,** 348–355 (2001).
34. Zheng, P., Wang, H., Bavister, B. D. & Ji, W. Maturation of rhesus monkey oocytes in chemically defined culture media and their functional assessment by IVF and embryo development. *Hum. Reprod.* **16,** 300–305 (2001).
35. Janovitz, T. *et al.* High-throughput sequencing reveals principles of adeno-associated virus serotype 2 integration. *J. Virol.* **87,** 8559–8568 (2013).
36. Cohn, L. B. *et al.* HIV-1 integration landscape during latent and active infection. *Cell* **160,** 420–432 (2015).
37. Xiao, J. *et al.* Rearrangement structure-independent strategy of CNV breakpoint analysis. *Mol. Genet. Genomics* **289,** 755–763 (2014).
38. Du, R. *et al.* Efficient typing of copy number variations in a segmental duplication-mediated rearrangement hotspot using multiplex competitive amplification. *J. Hum. Genet.* **57,** 545–551 (2012).
39. Kalin, N. H., Shelton, S. E., Davidson, R. J. & Kelley, A. E. The primate amygdala mediates acute fear but not the behavioral and physiological components of anxious temperament. *J. Neurosci.* **21,** 2067–2074 (2001).
40. Kalin, N. H., Shelton, S. E. & Davidson, R. J. The role of the central nucleus of the amygdala in mediating fear and anxiety in the primate. *J. Neurosci.* **24,** 5506–5515 (2004).
41. Rogers, J., Shelton, S. E., Shelledy, W., Garcia, R. & Kalin, N. H. Genetic influences on behavioral inhibition and anxiety in juvenile rhesus macaques. *Genes Brain Behav.* **7,** 463–469 (2008).
42. Wang, K. *et al.* MapSplice: accurate mapping of RNA-seq reads for splice junction discovery. *Nucleic Acids Res.* **38,** e178 (2010).
43. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11,** R106 (2010).
44. Liu, Z. *et al.* Generation of macaques with sperm derived from juvenile monkey testicular xenografts. *Cell Res.* **26,** 139–142 (2015).

© 2016 Macmillan Publishers Limited. All rights reserved



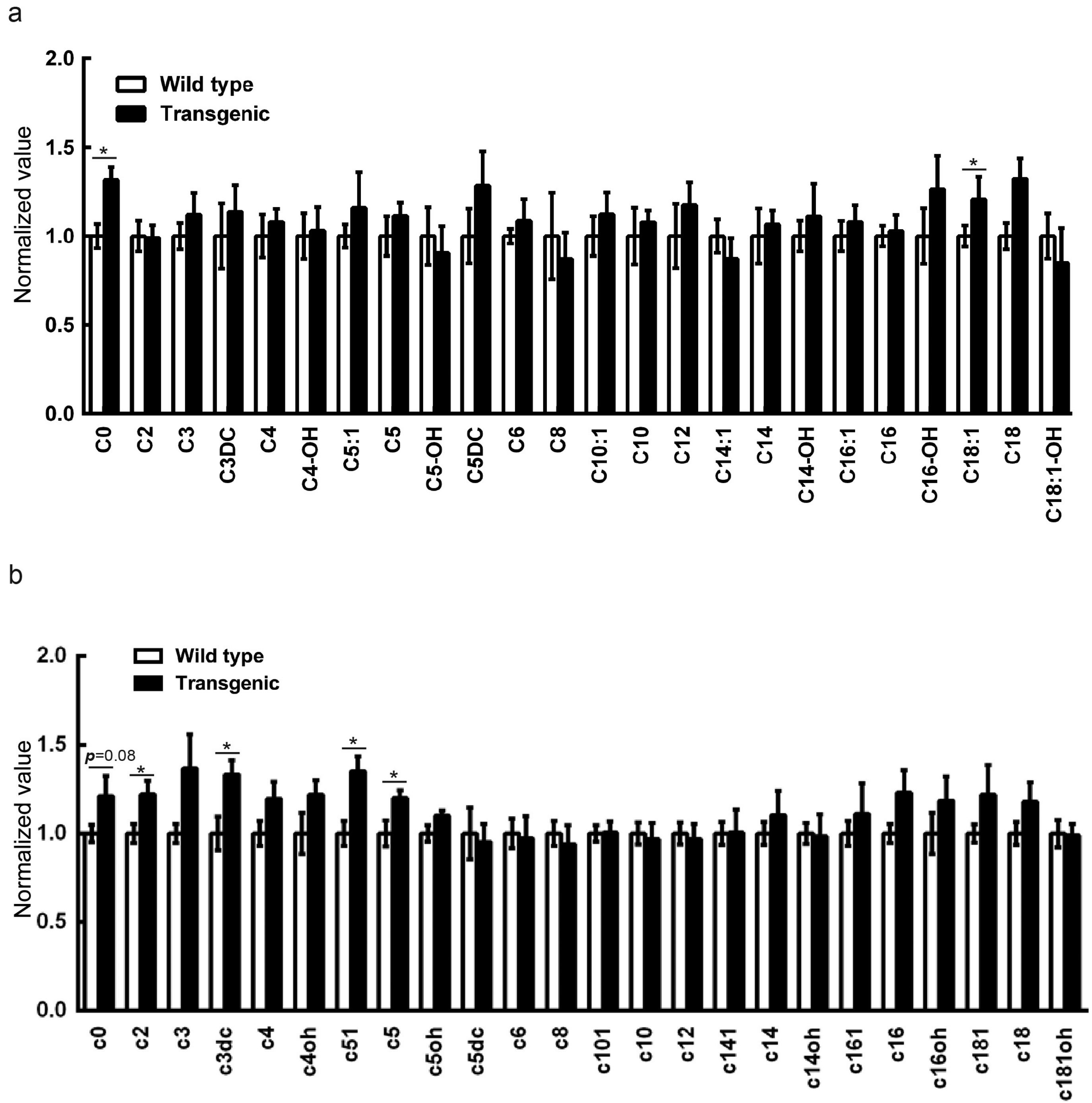
Extended Data Figure 1 | Analysis of genomic integration sites of transgenes in F_0 TG monkeys. a, Genomic DNA were extracted from hair roots of WT and TG monkeys, digested and analysed by agarose gel. Radioactive probe labelled by ^{32}P were prepared against the hMeCP2-2a-GFP transgene. Blots were transferred to membrane and hybridized with the probes. Images were acquired by exposing the blot to a phosphor-imager. Asterisk indicates target band. **b**, Flowchart of deep-sequencing-based methods for identifying genomic integration sites of lentiviral transgenes. Genomic integration sites of lentiviral transgenes are composed as three parts, endogenous genomic segments (black), LTRs (blue) and transgenes (red). Genomic DNA was sonicated, end-repaired and ligated to a Y-shaped adaptor, then subjected to two rounds of amplifications by the LTR1 + YP1 and LTR2 + YP2 primer sets. Illumina sequencing linkers were added onto segments and performed paired-end high-throughput sequencing. Target sequences containing LTR and endogenous genomic segments were collected and analysed. **c**, Comparison of copy numbers obtained from two methods among F_0 TG monkeys. Red denotes copy numbers from AccuCopy (*MECP2* and mCherry transgenes); blue denotes LTR insertion sites from deep-sequencing.

© 2016 Macmillan Publishers Limited. All rights reserved



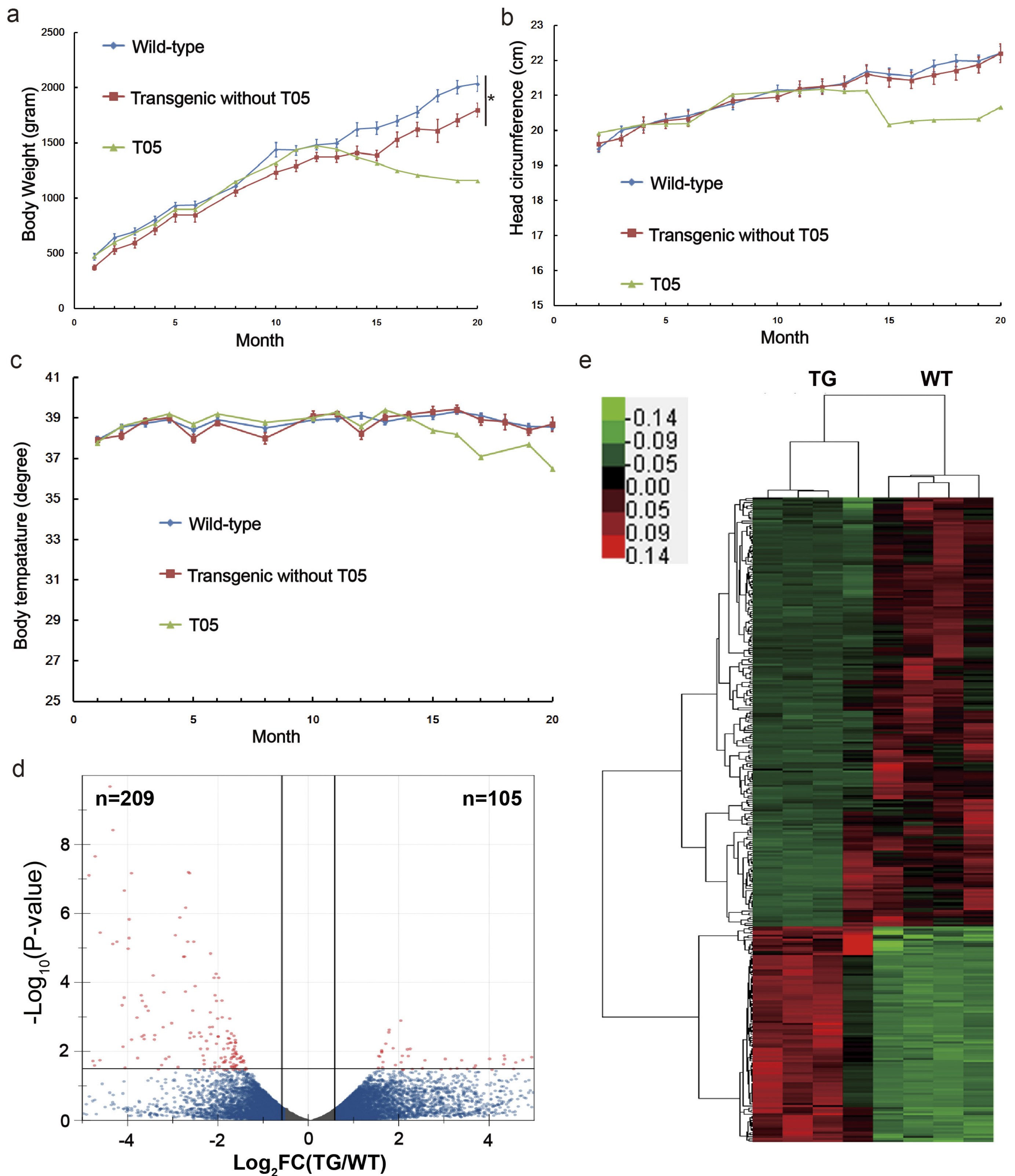
Extended Data Figure 2 | Physical parameters measured for monkeys. Developmental changes in the body weight of 8 TG and 18 WT monkeys. **a–g**, Body weight, abdominal circumference, heart rate, respiratory rate, head circumference, body temperature and head–truck length were measured for 8 *MECP2* TG and 18 WT monkeys. $*P < 0.05$ (Mann–Whitney U test). Error bars denote s.e.m.

© 2016 Macmillan Publishers Limited. All rights reserved



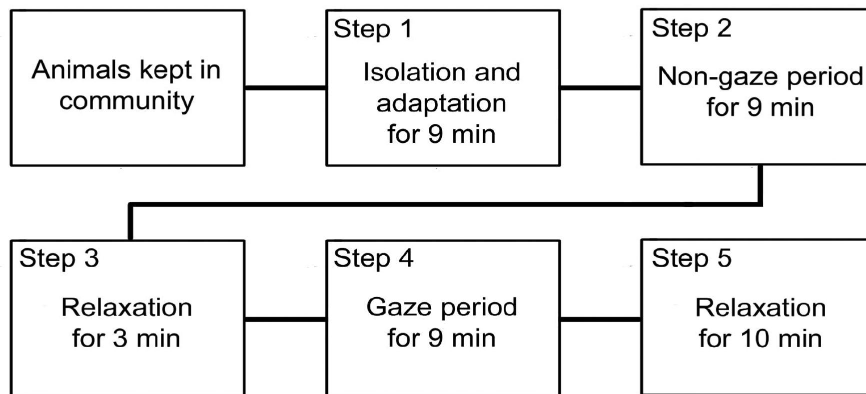
Extended Data Figure 3 | Fatty acid measurements for TG and WT monkeys. a, Blood samples collected at 18 months of age. **b**, Blood samples collected at 36 months of age. The blood levels of different forms of fatty acids were measured by mass spectrometry, with each bar represents results from three independent samples. C0, total fatty acid contents. All data are normalized to the average values of parallel blood samples from WT monkeys. $*P < 0.05$ (Student's *t*-test). Error bars denote s.e.m.

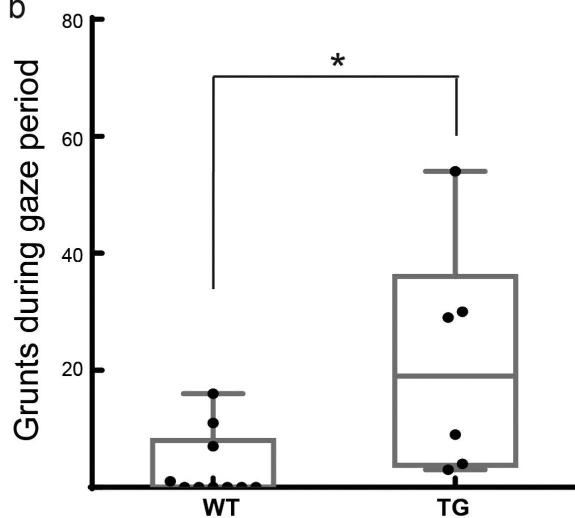
© 2016 Macmillan Publishers Limited. All rights reserved

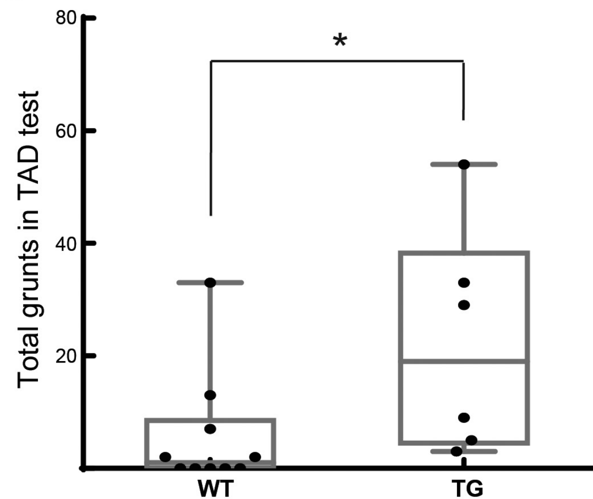


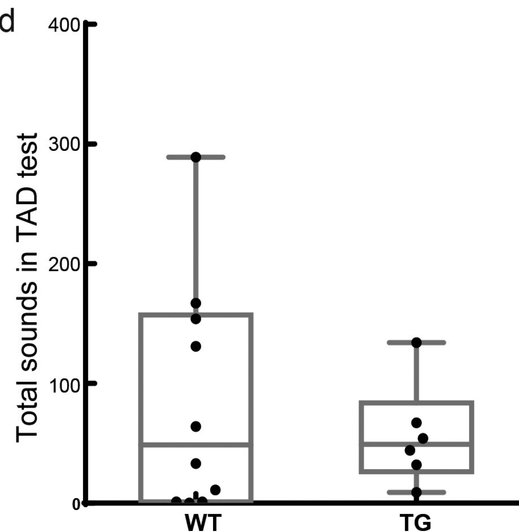
Extended Data Figure 4 | Physical growth parameters measured for monkey T05 and transcriptome analysis of *MECP2* transgenic monkey. **a**–**c**, Body weight, head circumference and body temperature were measured for monkey T05. *$P < 0.05$ (Mann–Whitney U test), together with the average data from all other TG and WT monkey monitored. The monkey T05 died at 20 months of age. **d**, Volcano map for alterations in gene expression in the TG monkeys (T14, T05, T07 and T09), as compared to four WT monkeys. Red dots denote genes with a >2-fold change (FC) in expression ($\log_2(FC(TG/WT)) > 1$ or < -1). Blue dots denote genes with no significant change in expression ($P > 0.05$). **e**, Heat map representation of the selected genes involved in metabolic pathways and brain development. Gene expression is coded in pseudocolour scale (−0.14 to 0.14). Red denotes higher expression in TG monkeys; green denotes lower expression in TG monkeys, as compared to WT monkeys. Error bars denote s.e.m.

© 2016 Macmillan Publishers Limited. All rights reserved



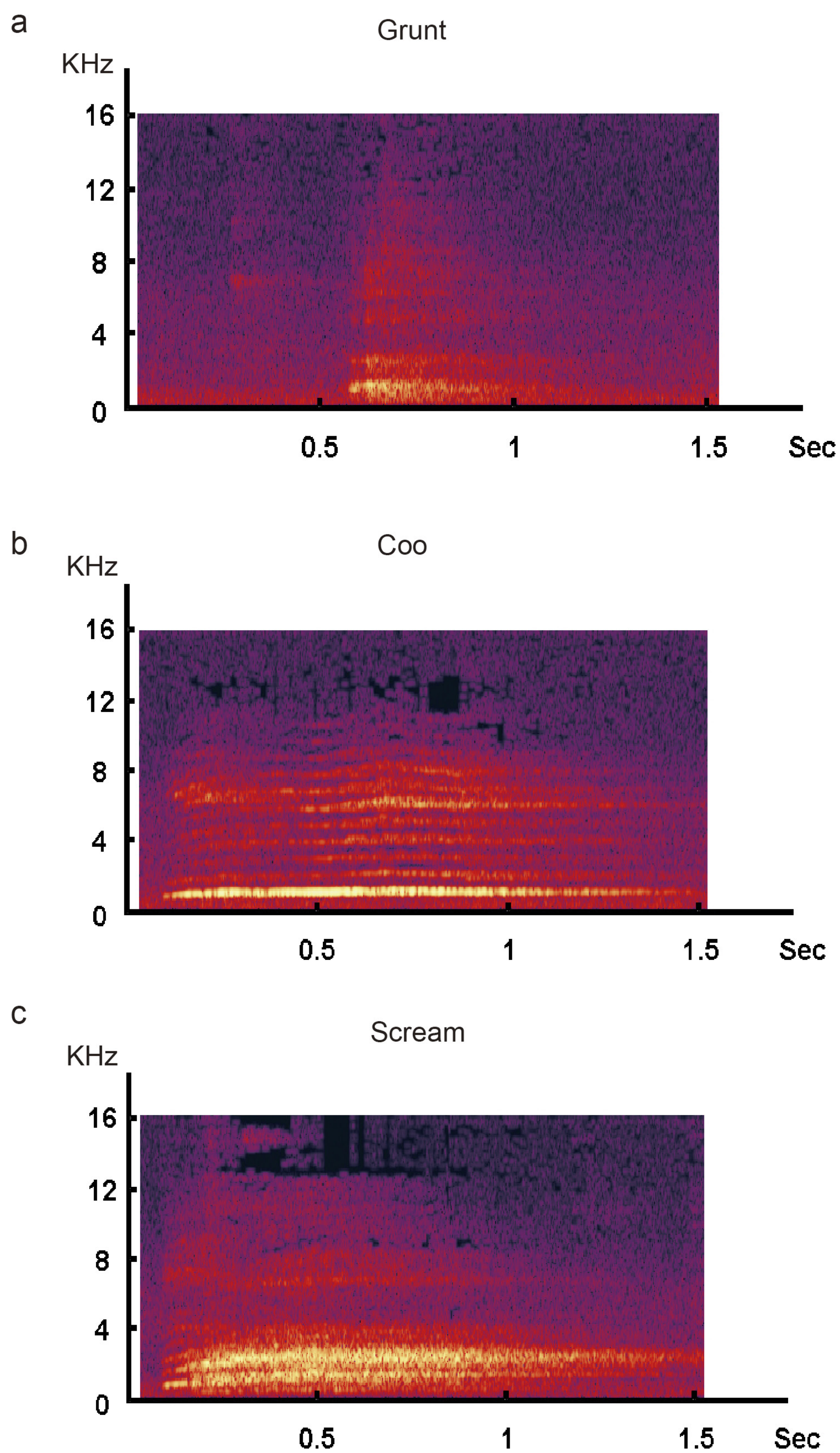






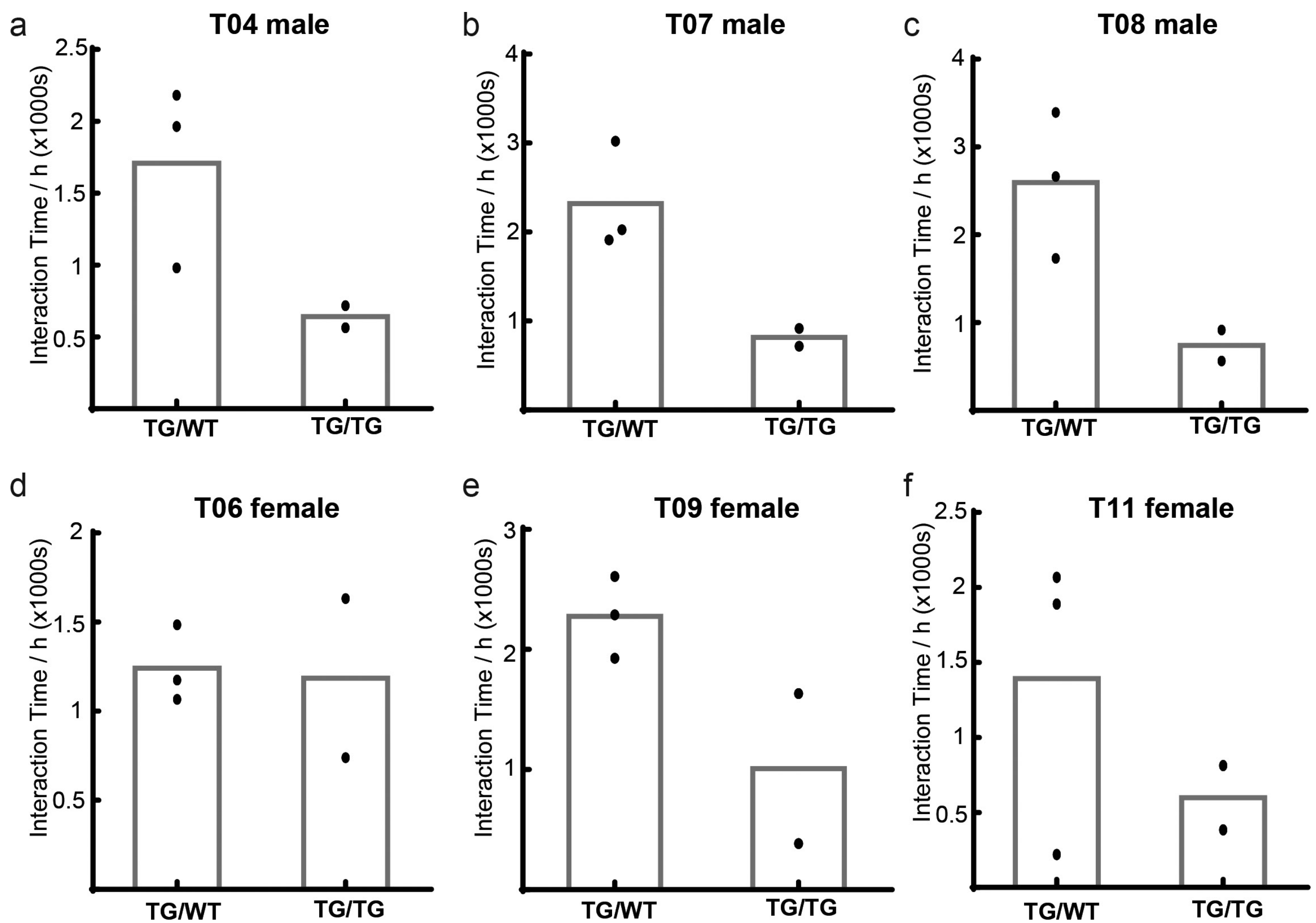
Extended Data Figure 5 | Anxiety responses in *MECP2* TG monkeys. **a**, Schematic illustration of the protocol of TAD test. **b**, Boxplots of the total numbers of grunt sounds made by WT and TG monkeys during the gaze period ('step 4' in the TAD test) at 36 months of age. **c**, Total grunts responses of wild-type and transgenic monkeys during TAD tests. **d**, Total vocal responses of wild-type and transgenic monkeys during TAD tests. $^{*}P < 0.05$ (Student's t-test). Ends of whiskers represent the minimum and maximum of data points. The line within box represents the median (odd numbers of data points) or second quartile (even number of data points). The bottom and top edge of the box represents the first and the third quartile, respectively.

© 2016 Macmillan Publishers Limited. All rights reserved



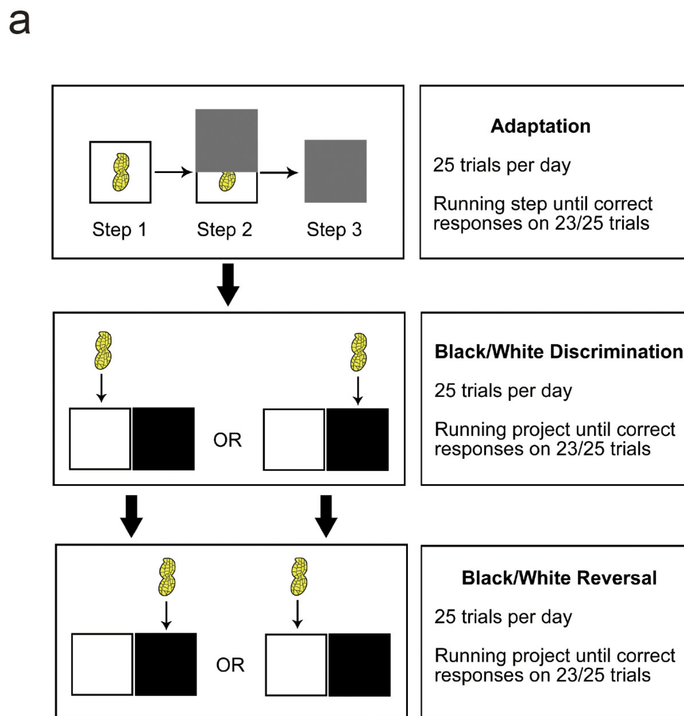
Extended Data Figure 6 | Spectrograms of typical sounds produced by the monkey in the TAD test. a–c, The power at different frequencies (ordinate) is colour-coded (red denotes higher power). Three typical sounds, grunt (**a**), coo (**b**) and scream (**c**), are shown.

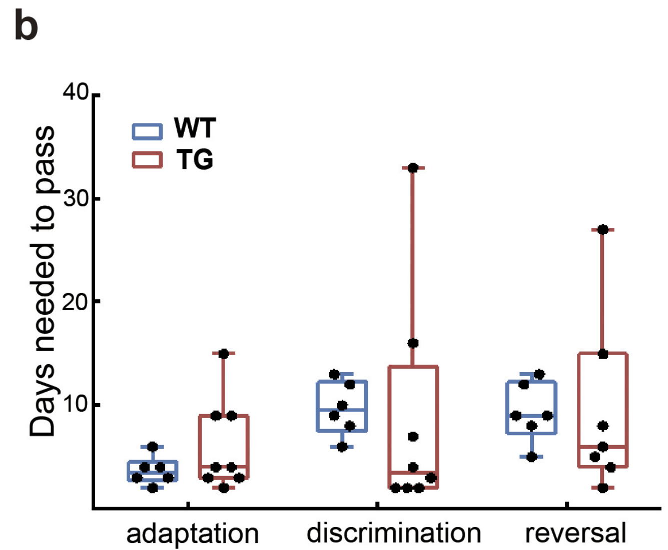
© 2016 Macmillan Publishers Limited. All rights reserved

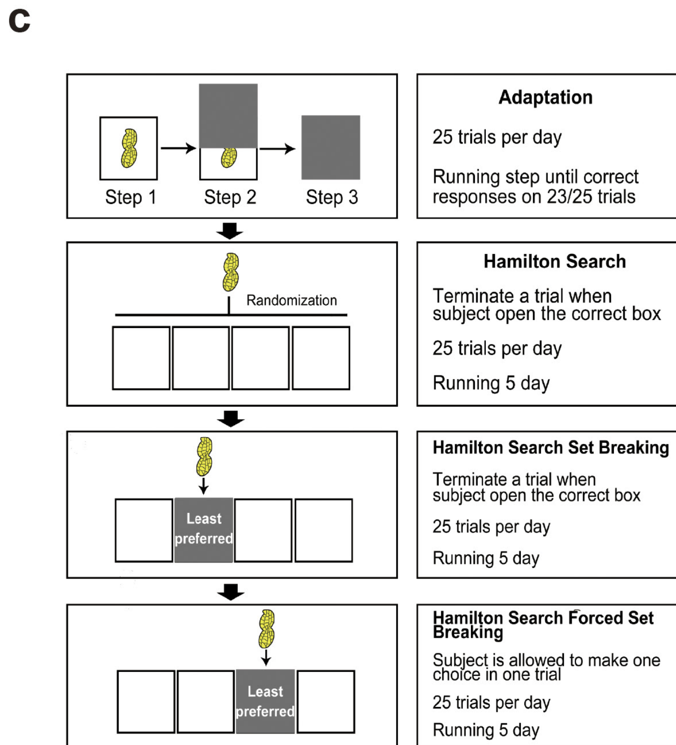


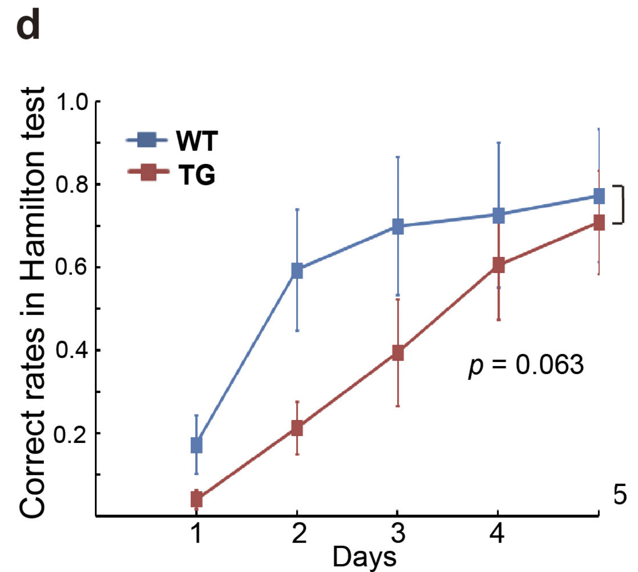
Extended Data Figure 7 | Social interaction between monkeys from the same group (familiar pairing). a–f, The average total time spent in sitting together during pairing in an isolated observation cage for each TG monkey (T04, T07, T08, T06, T09 and T11) with either a WT or a TG monkey was presented individually for six TG monkeys tested. T04 (**a**), T07 (**b**), T08 (**c**), T06 (**d**), T09 (**e**) and T11 (**f**). (See Supplementary Table 4c for grouping.) Each observation lasted 60 min daily for 5 days.

© 2016 Macmillan Publishers Limited. All rights reserved



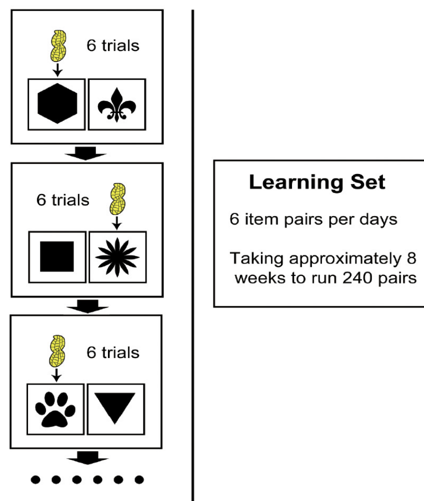


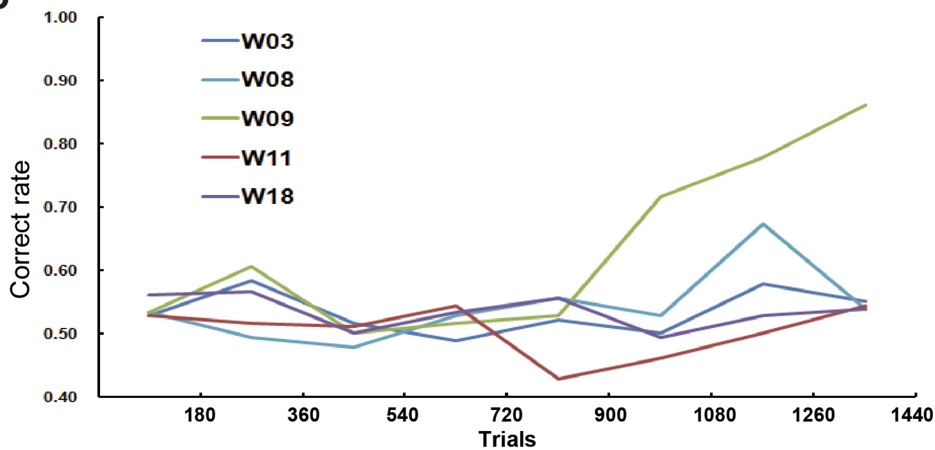


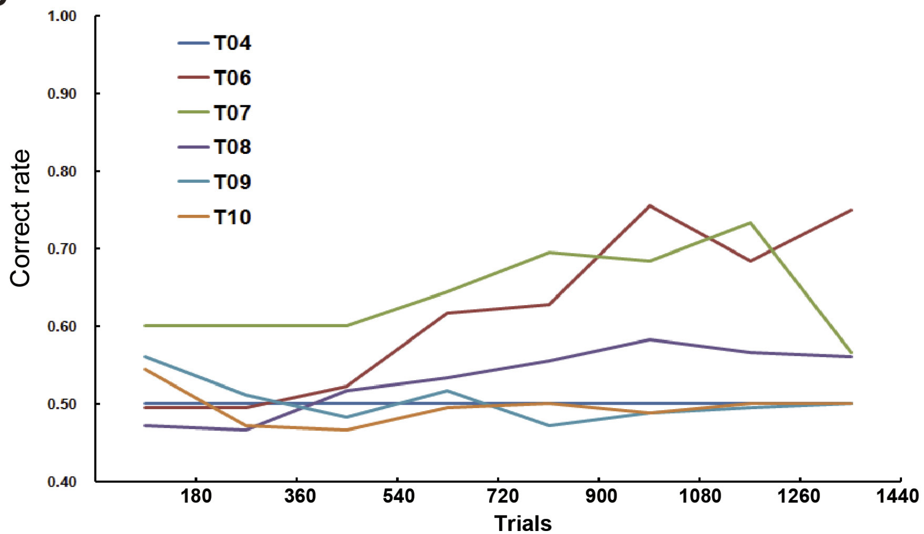


Extended Data Figure 8 | Schematic illustration of experimental procedures of WGTA test. a, Black/white test. **b**, Boxplots of days required to pass the adaptation, discrimination and reversal steps in the black/white test for six WT and eight TG monkeys. Ends of whiskers represent the minimum and maximum of data points. The line within box represents the median (odd numbers of data points) or second quartile (even number of data points). The bottom and top edge of the box represents the first and the third quartile, respectively. **c**, Hamilton search test. **d**, Learning curves for the Hamilton forced set-breaking test after passing the black/white test (for six WT and seven TG monkeys). The difference between the two groups was at a significance level of $P = 0.06$ (Mann–Whitney U test). Error bars denote s.e.m.

© 2016 Macmillan Publishers Limited. All rights reserved







Extended Data Figure 9 | Performance of WT and TG monkeys in learning set of WGTA test. a, Learning set test. Correct rate of monkeys in the reward-shape association learning test plotted individually against trials, with data points represents average correct rates over 180 trials. **b**, WT monkeys. **c**, TG monkeys.

© 2016 Macmillan Publishers Limited. All rights reserved

Extended Data Table 1 | Summary of *MECP2* transgenic monkey cohorts

a.

Animal code	Date of birth (y/m/d)	Gender	Transgenes	Copy numbers identified by AccuCopy		Note
				hMECP2	mCherry	
T01	2011/2/28	F	*HA-hMeCP2-2a-GFP+mCherry*	3.4	N/A	Stillbirth
T02	2011/5/30	F	*HA-hMeCP2-2a-GFP+mCherry*	1.0	N/A	Stillbirth
T03	2011/5/30	F	*HA-hMeCP2-2a-GFP+mCherry*	N/A	N/A	Stillbirth
T04	2011/4/9	M	*HA-hMeCP2-2a-GFP+mCherry*	1.0	1.9	Live
T05	2011/3/18	F	*HA-hMeCP2-2a-GFP+mCherry*	1.2	3.8	Live, deceased at 2013/3/12
T06	2011/4/9	F	*HA-hMeCP2-2a-GFP+mCherry*	7.3	6	Live
T07	2011/4/10	M	*HA-hMeCP2-2a-GFP+mCherry*	5.7	4.3	Live, deceased at 2015/2/4
T08	2011/4/12	M	*HA-hMeCP2-2a-GFP+mCherry*	2.9	2.3	Live
T09	2011/4/16	F	*HA-hMeCP2-2a-GFP+mCherry*	2.4	3	Live, deceased at 2014/11/8
T10	2011/4/16	F	*HA-hMeCP2-2a-GFP+mCherry*	1.1	2	Live
T11	2011/6/18	F	*HA-hMeCP2-2a-GFP+mCherry*	1.9	1.3	Live
T12	2011/4/9	F	*mCherry*	0	N/A	Stillbirth

b.

Animal code	Date of birth (y/m/d)	Gender	Transgenes	Copy numbers identified by AccuCopy		Note
				hMECP2	mCherry	
TF1-1	2014/11/13	M	*HA-hMeCP2-2a-GFP+mCherry*	2.5	0.6	Deceased 3 days after birth
TF1-2	2014/11/18	M	*HA-hMeCP2-2a-GFP+mCherry*	2.5	0.5	Live
TF1-3	2014/12/7	F	*HA-hMeCP2-2a-GFP*	5.5	0	Live
TF1-4	2014/11/27	M	*HA-hMeCP2-2a-GFP+mCherry*	3.3	0.9	Live
TF1-5	2014/11/27	F	*HA-hMeCP2-2a-GFP+mCherry*	6.0	1.0	Live

a, Summary of *MECP2* F_0 transgenic monkeys. **b**, Summary of *MECP2* F_1 transgenic monkeys.

© 2016 Macmillan Publishers Limited. All rights reserved