Coding of facial expressions of pain in the laboratory mouse

Dale J Langford[1], Andrea L Bailey[1], Mona Lisa Chanda[1], Sarah E Clarke[1], Tanya E Drummond[1], Stephanie Echols[2], Sarah Glick[1], Joelle Ingrao[1], Tammy Klassen-Ross[2], Michael L LaCroix-Fralish[1], Lynn Matsumiya[1], Robert E Sorge[1], Susana G Sotocinal[1], John M Tabaka[1], David Wong[2], Arn M J M van den Maagdenberg[3,4], Michel D Ferrari[4], Kenneth D Craig[2] & Jeffrey S Mogil[1]

Facial expression is widely used as a measure of pain in infants; whether nonhuman animals display such pain expressions has never been systematically assessed. We developed the mouse grimace scale (MGS), a standardized behavioral coding system with high accuracy and reliability; assays involving noxious stimuli of moderate duration are accompanied by facial expressions of pain. This measure of spontaneously emitted pain may provide insight into the subjective pain experience of mice.

© 2010 Nature America, Inc. All rights reserved.

Darwin famously asserted that nonhuman animals are capable of expressing emotion (including pain) through facial expression[1], and that such expression may be both innate and adaptive. Observations that a similar facial expression of pain is displayed by neonates and even by the congenitally blind lend support to this notion[2]. The ability to communicate one's pain experience to others may benefit both the sender and receiver, such that help might be offered or a warning signal heeded[3]. However, emitting visible signals of pain might also be selected against in prey animals like rodents.

Facial expressions of virtually every common emotion, including pain, have been well characterized in humans, and can be reliably coded using the anatomically based 'action units' of the facial action coding system[4]. Similar scales have been adapted and have become useful tools in the assessment of pain and analgesia in clinical populations in which verbal communication is limited or nonexistent, such as infants and those with cognitive impairments[2].

Despite evidence that nonhuman mammals including rats[5] exhibit facial expressions of other emotional states, there has been no study of facial expressions of pain in any nonhuman species. Considering the pain field's heavy and continuing dependence on rodent models and the paucity of usable measures of spontaneous (as opposed to experimenter-evoked) pain in animals[6], the ability to reliably and accurately detect pain, in real time, using facial expression might offer a unique and powerful scientific tool in addition to having obvious benefits for veterinary medicine.

To develop the scale, we used a classical preclinical pain assay, the 0.9% acetic acid abdominal constriction test (Online Methods). We placed mice singly in plexiglas observation cubicles (9 × 5 × 5 cm high) with digital video cameras positioned at either end and filmed them for 30 min before and after administering the painful stimulus. We grabbed frames from video taken before injection at ~3-min intervals and taken after injection as mice experienced an abdominal constriction. We cropped the resultant JPEG files such that body position was no longer visible.

We compiled collages of 'no-pain' (baseline) and 'pain' photographs, along with individual no-pain and pain photos for each mouse, to devise a coding system, the mouse grimace scale (MGS), consisting of five facial features (action units) perceived by human facial pain expression experts (members of the K.D.C. laboratory) as potentially reliable indices of pain (**Fig. 1**). We defined action units as follows. Orbital tightening is narrowing of the orbital area, with a tightly closed eyelid or an eye squeeze (denoted by wrinkle around eye). Nose bulge is a rounded extension of skin visible on the bridge of the nose. Cheek bulge refers to convex appearance of the cheek muscle (between eye and whiskers) from its baseline position. Ear position refers to ears pulled apart and back from their baseline position or featuring vertical ridges that form owing to tips of ears being drawn back. Whisker change is movement of whiskers from their baseline position either backward, against the face or forward, as if standing on end; whiskers may also clump together. We observed obvious piloerection in most cases (data not shown). Note that three of these features are identical to those observed in humans: orbital tightening, nose bulge and cheek bulge[7], supporting Darwin's century-old prediction that facial expressions are evolutionarily conserved.

After a brief, formal training session, we presented randomized sets of digital photos to seven coders blinded to pain status to assess accuracy and reliability of the MGS. Coders assigned a value of 0 (not present), 1 (moderately visible) or 2 (severe) for each action unit in each photo and made a global pain or no-pain judgment for each mouse. The average global accuracy was 72% and 81% for a coder with one year of experience (**Supplementary Fig. 1a**). Of the errors, false alarms were rarer than misses. The scale displayed high internal consistency (Cronbach's $\alpha = 0.89$), as well as high inter-rater reliability assessed by intraclass correlation coefficient ($ICC_{average} = 0.90$; **Supplementary Fig. 1b**). Coders consistently assigned higher overall scores (average of all five action units) to pain versus no-pain photographs ($P < 0.001$). Later use of high-definition video cameras yielded substantial

[1]Department of Psychology and Alan Edwards Centre for Research on Pain, McGill University, Montreal, Quebec, Canada. [2]Department of Psychology, University of British Columbia, Vancouver, British Columbia, Canada. [3]Department of Human Genetics, Leiden University Medical Centre, Leiden, The Netherlands. [4]Department of Neurology, Leiden University Medical Centre, Leiden, The Netherlands. Correspondence should be addressed to J.S.M. (jeffrey.mogil@mcgill.ca).

RECEIVED 18 FEBRUARY; ACCEPTED 31 MARCH; PUBLISHED ONLINE 9 MAY 2010; DOI:10.1038/NMETH.1455

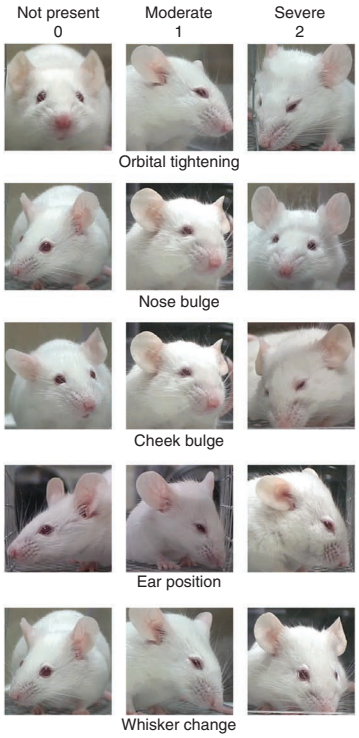

Figure 1 | In the MGS, intensity of each feature is coded on a three-point scale. For each of the five features, images of mice exhibiting behavior corresponding to the three values are shown.

improvement in accuracy, such that over 97% of pain versus no-pain calls were correct (**Supplementary Fig. 1a**).

We also assessed the scale on 14 commonly used preclinical pain assays (Online Methods), comparing average action unit scores assigned to 'pain' photos to those assigned to 'no-pain' photos and then comparing the resulting difference scores to zero by one-sample *t*-tests. The MGS revealed significant changes from baseline in some assays but not in others (**Fig. 2a**). Noxious stimuli of moderate duration (10 min–4 h) were most likely to be associated with a 'pain face'. Noxious stimuli applied to deep tissues (joints and viscera) yielded higher MGS difference scores than when applied more superficially (**Fig. 2a**,**b**). This specificity may be due to an inability to suppress painful facial expressions for stimuli of longer duration and/or internal origin because inhibition of such expression would otherwise be adaptive in reducing vulnerability to potential predators or aggressive conspecifics. The presumed neuropathic pain after experimental nerve injury is of very long duration, however, and was not associated with a pain face. This may be because of its paroxysmal nature (and our inability to recognize an ongoing paroxysm), but no pain face was induced by the application of a binder clip (700 g) to the ipsilateral hindpaw of nerve-damaged mice (data not shown). Alternatively, such specificity may reflect differences in the strength of the affective component of prolonged and/or deep (especially visceral) pain, as reported in humans[8]. Chemical lesioning of two structures previously associated with pain affect in the rodent, the anterior cingulate cortex and the amygdala[9], did not affect MGS scores (data not shown). However, ibotenic acid lesions of the rostral anterior insula, which is activated in humans by pain with an important emotional component[10], produced attenuation of the pain face (**Fig. 2c**) without affecting abdominal constriction behavior (**Fig. 2d**). Individuals with insular lesions can present with pain asymbolia, in which emotional responses to pain are markedly reduced but pain thresholds are unaffected[11], and thus it is tempting to speculate that the pain face reflects pain-related negative affect. Finally, as Darwin[1] and others[12] more recently have suggested, such facial expression may serve an egocentric sensory function by modifying sensory input in a beneficial direction; in this case, to limit sensory input thereby potentially reducing pain perception. Additional research is necessary to choose among these alternatives, and rodents might be useful models to do so, given that mechanisms whereby pain experience is encoded by facial expression in humans are still not understood[2].

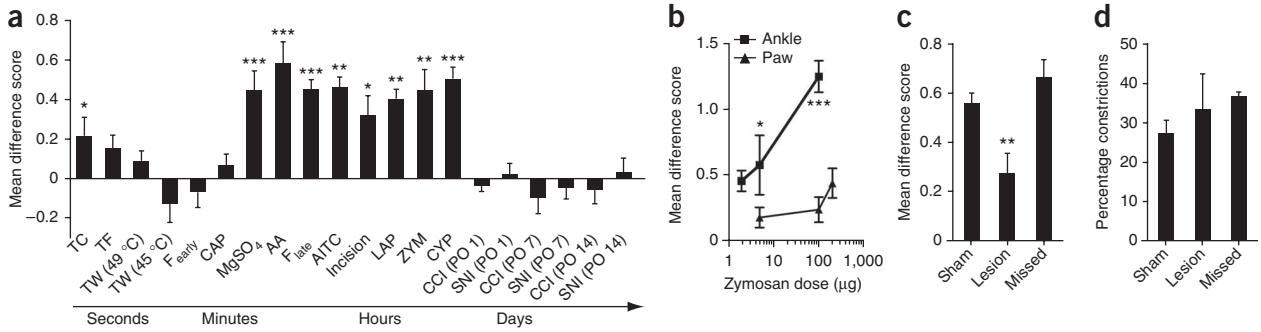

Figure 2 | Specificity of the MGS. (**a**) Nociceptive assays of moderate duration (~10 min to ~12 h) feature a 'pain face'. Assays are listed from left to right by (approximate) increasing duration of known behavioral measures of nociception and/or hypersensitivity. Error bars, s.e.m.; TC, 700 g tail clip ($n = 20$); TF, tail flick away from radiant heat ($n = 20$); TW, tail withdrawal from hot water ($n = 11$); F_{early}, 5% formalin test, early phase (0–5 min) ($n = 12$); CAP, 12.5% intraplantar capsaicin ($n = 8$); $MgSO_4$, 120 mg kg^{-1} $MgSO_4$-induced abdominal constriction ($n = 10$); AA, 0.9% acetic acid–induced abdominal constriction ($n = 14$); F_{late}, 5% formalin test, late phase (15–60 min) ($n = 11$); AITC, 5% intraplantar allyl isothiocyanate (mustard oil) ($n = 8$); incision, hind paw incision ($n = 8$); LAP, laparotomy ($n = 12$); ZYM, 200 µg intraplantar zymosan ($n = 8$); CYP, 400 mg kg^{-1} cyclophosphamide-induced bladder cystitis ($n = 12$); CCI, chronic constriction injury ($n = 8$); SNI, spared nerve injury ($n = 9$); PO, postoperative day. Difference scores were compared to zero by one-way Student's *t*-tests; $^*P < 0.05$, $^{**}P < 0.01$ and $^{***}P < 0.001$ (all uncorrected). All significant findings in **a** remained so after correction for multiple comparisons, except for TC. Time labels indicate the approximate duration of the assays. (**b**) Stimulus intensity dependence of MGS difference scores after injection of zymosan into the plantar hindpaw or ankle joint. Error bars, s.e.m.; $n = 6$ (all ankle doses), $n = 8$ (paw 200 µg); $n = 10$ (paw 100 µg); $n = 15$ (paw 5 µg). The deeper (ankle) injection resulted in a large shift of the dose-response curve. $^*P < 0.05$, $^{***}P < 0.001$ compared to other group (at some dose) by two-way Student's *t*-test. (**c**,**d**) Mean difference scores (**c**) and mean percentage of samples featuring abdominal constrictions (**d**) of mice receiving saline lesions (sham; $n = 6$), correctly targeted ibotenic acid lesions of the rostral anterior insula (lesion; $n = 4$) and incorrectly targeted ibotenic acid lesions (missed; $n = 2$). Error bars, s.e.m.; $^{**}P < 0.01$ compared to other groups (one-way ANOVA: $F_{2,9} = 7.9$).

© 2010 Nature America, Inc. All rights reserved.

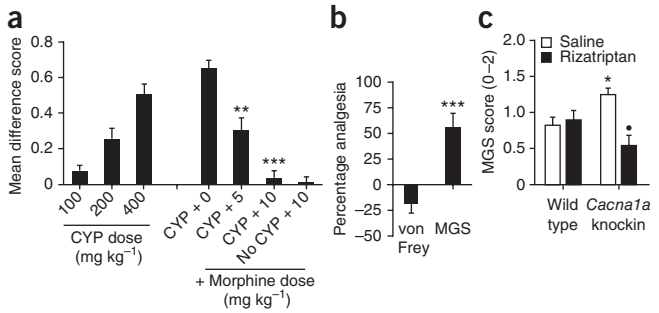

Figure 3 | Validity and utility of the MGS. (**a**) Correlation between cyclophosphamide (CYP) dose and MGS difference scores (main effect of CYP dose: $F_{2,17} = 8.9$, $P < 0.005$) and dose-dependent morphine reversal of 400 mg kg^{-1} CYP–induced pain face (main effect of morphine dose: $F_{2,12} = 26.6$, $P < 0.001$). Error bars, s.e.m.; $n = 4$ (100, 200, CYP + 10, No CYP + 10), $n = 5$ (CYP + 5), $n = 6$ (CYP + 0), $n = 12$ (400); $^{**}P < 0.01$; $^{***}P < 0.001$ compared to CYP plus saline (CYP + 0) by Dunnett's test. Note that administration of 10 mg kg^{-1} morphine (no CYP + 10) produces no change from baseline. (**b**) Analgesic efficacy of 300 mg kg^{-1} acetaminophen after administration of 100 µg zymosan as assayed by MGS but not by von Frey fiber-measured mechanical allodynia. Error bars, s.e.m.; $n = 4$ (MGS), $n = 8$ (von Frey); $^{***}P < 0.001$ compared to zero (or other group) by *t*-test. (**c**) "Baseline" MGS scores in restrained transgenic *Cacna1a* knockin mutant mice bearing the human familial hemiplegic migraine, type 1 R192Q missense mutation in the encoded protein compared to wild-type controls treated with saline or the antimigraine drug, rizatriptan (50 mg kg^{-1}, intraperitoneal). Error bars, s.e.m.; $n = 4$ (rizatriptan), $n = 6$ (saline). After two-way ANOVA revealed a significant genotype × drug interaction ($F_{1,16} = 10.9$; $P = 0.005$), *t*-tests were performed. $^{*}P < 0.05$ compared to wild type; $^{\bullet}P < 0.05$ compared to saline within genotype.

By practical necessity, we conducted tests of superficial pain requiring precise aiming of a noxious stimulus using a restraining apparatus that may have induced restraint stress, and this elevated baseline MGS scores; however, in superficial assays laboriously conducted using standard cubicles, difference scores (Online Methods) were nonsignificant ($P > 0.05$; data not shown). Aside from stress, other states such as sleep, grooming and illness have some (but not all) facial features in common with pain (**Supplementary Fig. 2**), and we advise that such photographs not be coded.

We validated the scale first by determining that pain scores on the MGS were related to stimulus intensity. Because we found scores to be unvarying with stimulus intensity during the exhibition of observable pain behavior in the abdominal constriction test (data not shown), we used two assays devoid of obvious behavioral paroxysms, intraplantar or intraarticular zymosan and cyclophosphamide-induced bladder cystitis, and altered stimulus intensity by varying the injected dose. We found positive and linear relationships between dose and MGS scores (**Figs. 2b** and **3a**). We also observed a dose-dependent reversal of the cyclophosphamide-induced pain face upon administration of morphine sulfate (5 and 10 mg kg^{-1}), in the absence of any effect on facial expression of morphine itself (**Fig. 3a**).

We tested the sensitivity and utility of the MGS in two additional experiments. Zymosan inflammation in mice produced no overt spontaneous behaviors but was associated with robust mechanical hypersensitivity (allodynia) and a pain face. The weak analgesic, acetaminophen (300 mg kg^{-1}), produced no detectable antiallodynic effects in our hands but reduced MGS difference scores by more than 50% (**Fig. 3b** and **Supplementary Fig. 3**). As part of an ongoing behavioral characterization of transgenic *Cacna1a* knockin mutant mice bearing the human familial hemiplegic migraine, type 1 R192Q missense mutation in the encoded protein that causes gain-of-function effects on neuronal voltage-gated $Ca_V2.1$ calcium channels[13], we observed a 'baseline' pain face in restrained mutant mice compared to wild-type mice, which was prevented by 50 mg kg^{-1} rizatriptan (**Fig. 3c**), suggestive of endogenous pain in the mutant mice.

In conclusion, we developed a standardized facial coding system, the MGS, characterized by high accuracy, reliability and validity, which can be used to assess pain in the laboratory mouse for veterinary care and drug development. A detailed training manual will be provided to interested parties upon request. The MGS has the advantage of measuring a (likely involuntary) behavior spontaneously emitted by the subject itself, in the absence of a looming experimenter and under conditions perfectly blinded to pain status. The demonstration of a 'pain face' in the mouse would seem to support evolutionary accounts of human facial pain expression[2] over popular operant explanations. Finally, given that mice can be affected by the pain status of a familiar conspecific[14] and that female mice prefer to maintain close proximity to familiars in pain[15], it is also conceivable that this social modulation is mediated by attending to the facial expression of other mice.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturemethods/.

Note: Supplementary information is available on the Nature Methods website.

ACKNOWLEDGMENTS

This work was funded by the Louise and Alan Edwards Foundation (J.S.M.). D.J.L. was supported by a Canada Graduate Scholarship doctoral award from the Canadian Institutes of Health Research.

AUTHOR CONTRIBUTIONS

D.J.L. and J.S.M. conceived the experiments, and D.J.L. and S.G.S. were primarily responsible for their execution. A.L.B., M.L.C., S.E.C., T.E.D., S.G., M.L.L.-F., L.M., R.E.S. and J.M.T. participated in data collection. S.E., J.I., T.K.-R., D.W. and K.D.C. were responsible for designing the scale. A.M.J.M.v.d.M. and M.D.F. provided transgenic knock-in mice and edited portions of the manuscript. J.S.M. and D.J.L. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturemethods/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Darwin, C. *The Expression of the Emotions in Man and Animals* (Albemarle, 1872).
2. Williams, A.C. *Behav. Brain Sci.* **25**, 439–455 (2002).
3. Craig, K.D. *Can. Psychol.* **50**, 22–32 (2009).
4. Ekman, P. & Friesen, W. *Facial Action Coding System* (Consulting Psychologists Press, 1978).
5. Grill, H.J. & Norgren, R. *Brain Res.* **143**, 263–279 (1978).
6. Mogil, J.S. *Nat. Rev. Neurosci.* **10**, 283–294 (2009).
7. Prkachin, K.M. *Pain* **51**, 297–306 (1992).
8. Strigo, I.A., Bushnell, M.C., Boivin, M. & Duncan, G.H. *Pain* **97**, 235–246 (2002).
9. Gao, Y.J., Ren, W.H., Zhang, Y.Q. & Zhao, Z.Q. *Pain* **110**, 343–353 (2004).
10. Schweinhardt, P. *et al. Neuroimage* **32**, 256–265 (2006).
11. Berthier, M., Starkstein, S. & Leiguarda, R. *Cortex* **23**, 673–678 (1987).
12. Susskind, J.M. *et al. Nat. Neurosci.* **11**, 843–850 (2008).
13. van den Maagdenberg, A.M.J.M. *et al. Neuron* **41**, 701–710 (2004).
14. Langford, D.J. *et al. Science* **312**, 1967–1970 (2006).
15. Langford, D.J. *et al. Soc. Neurosci.* (in the press).

© 2010 Nature America, Inc. All rights reserved.

ONLINE METHODS

Animals. All mice (except in the 'migraine mouse' experiment) were CD-1 (ICR:Crl) mice, aged 6–18 weeks, bred in our vivarium with mice obtained from Charles River Laboratories. Mice were housed in groups of two or more, under a 14-h light, 10-h dark cycle (lights on at 07:00 h) in a temperature-controlled environment (20 ± 1 °C) with *ad lib* access to food (Harlan Teklad) and tap water. For each nociceptive assay a new cohort of mice was used, such that no mouse participated in more than one assay. Equal numbers of male and female mice were tested in each cohort. No statistically significant sex differences in baseline or poststimulus average MGS score were observed in any assay, although trends toward higher poststimulus MGS scores in female mice were observed in several cases. All animal experiments were approved by the McGill University (Downtown) animal care and use committee.

Compounds. All noxious chemical compounds and acetaminophen were obtained from Sigma-Aldrich and dissolved in physiological saline except when noted otherwise. Morphine sulfate was generously provided by Sandoz Canada.

Initial frame capture and scale development. Mice were individually placed on a table top in cubicles (9 × 5 × 5 cm) with two walls of transparent plexiglas and two side walls of removable stainless steel. Two digital video cameras were placed immediately outside both plexiglas walls to maximize the opportunity for clear headshots. Mice were acclimated and digitally videotaped (640 × 480 pixels) with black-and-white digital video cameras (Sony Handycam Digital Video Camera Recorder model DCR-SR100) for 30 min before injection (baseline; 'no-pain') and for 30 min after intraperitoneal injection of 0.9% acetic acid (after injection; 'pain'). Using Microsoft Windows Media Player, individual frames of the resultant WMV files were 'grabbed' whenever a clear, unobstructed head shot was observed. In the period after injection, frames were grabbed specifically during the exhibition of the writhing or stretching behavior (lengthwise constrictions of the abdominal musculature) normally used as the dependent measure in this assay. The resultant JPG files were cropped (so that body position was no longer visible) and autoadjusted for contrast and brightness in Adobe Photoshop CS 8.0. Higher-resolution (1,920 × 1,080 pixels) color videos (used in more recently performed experiments shown in **Figures 2b**,**c** and **3b** and **Supplementary Figure 1a**) were obtained using the Sony High Definition Handycam camcorder (model HDR-CX100).

Multiple collages of pain and no-pain photographs were compiled and sent to S.E., T.K., D.W. and K.D.C., who used these collages along with individual baseline and postinjection photos for each mouse tested to devise a coding system (MGS) consisting of facial action units (AUs) they perceived as potentially reliable indices of pain.

Accuracy and reliability determination. A detailed handout as well as a Microsoft PowerPoint presentation were prepared and distributed (by D.J.L.) to McGill graduate and undergraduate members of the J.S.M. lab. During a formal 1-h training session, each feature was discussed, and a prototypic expression for each intensity score (0–2) was displayed. Seven graduate and undergraduate student coders were then given ten randomized sets of 64 unlabeled photos to assess accuracy and reliability of the MGS. Accuracy was determined by global 'pain' versus 'no-pain' judgments; reliability by comparing average AU scores across coders, using the ICC (two-way mixed design; mean score across 10 orders reported)[16].

Nociceptive assays. The following assays were used (n = 8–20 mice per assay in **Figure 2a**; see below for sample sizes in other experiments), described below in alphabetical order by abbreviation. Detailed descriptions of most of these assays were reported previously[17,18]. All stimuli are thought to be noxious because mice either: (i) reflexively withdraw from them, (ii) exhibit obvious nocifensive or recuperative behaviors such as licking, biting or shaking in response to them or (iii) display hypersensitivity (hyperalgesia and/or allodynia) to evoking thermal and/or mechanical stimuli after the insult. In many cases, identical or highly similar insults are known to be painful in humans. But there is no direct evidence for the existence of spontaneous pain in many of these assays. In all cases mice were habituated for 30 min before testing began.

For the acetic acid abdominal constriction test (AA), 10 ml kg^{-1} of 0.9% acetic acid was injected intraperitoneally. Mice were immediately returned to their cubicles and filmed for 30 min after injection. For the allyl isothiocyanate (AITC) test, 20 µl of 5% AITC was injected subcutaneously into the plantar surface of the right hindpaw. Mice were immediately returned to their cubicles and filmed for 30 min after injection. For the capsaicin (CAP) test, 20 µl of 125 µg ml^{-1} capsaicin (dissolved in 80% saline, 10% Tween and 10% ethanol) was injected subcutaneously into the plantar surface of the right hindpaw. Mice were immediately returned to their cubicles and filmed for 20 min after injection. For the chronic constriction injury (CCI), surgery was performed under general anesthesia essentially as previously described[19]. Mice were returned to their home cages until testing. On test days (1, 7 and 14 d postoperative), mice were habituated for 30 min before being filmed for 1 h. For the cyclophosphamide (CYP) cystitis test, 10 ml kg^{-1} of CYP (100, 200 or 400 mg kg^{-1}) was injected intraperitoneally. This drug is converted to acrolein in the liver and collects in the bladder, causing a painful cystitis in humans and behavioral changes and allodynia referred to the lower abdomen in mice[20]. Mice remained in their home cages for 2 h and were then returned to their observation cubicles for 60 min before being filmed for 1 h.

For the formalin test (F_{early}, F_{late}), 20 µl of 5% formalin was injected subcutaneously to the plantar surface of the right hindpaw. Mice were returned to their cubicles and filmed for 60 min after injection. The early (acute; F_{early}) phase of the formalin test was conservatively defined as 0–5 min after injection and the late (tonic; F_{late}) phase as 15–60 min after injection. For the incision model, the hindpaw incision model of postoperative pain was performed under isoflurane-oxygen anesthesia as previously described[21] for the mouse, an adaptation of the rat model[22]. Once recovered from the anesthesia, mice were returned to their cubicles and filmed for 2 h.

For the laparotomy (LAP) model, a laparotomy, designed to mimic a sham ventral ovariectomy, was performed under isoflurane-oxygen anesthesia. After shaving and disinfection, a 2-cm midline incision was made using iris scissors. Muscle layers were closed with polydioxanone suture 5-0 and skin edges apposed

© 2010 Nature America, Inc. All rights reserved.

doi:10.1038/nmeth.1455

using tissue glue. Once recovered from anesthesia, mice were returned to their cubicles and filmed for 2 h. For the magnesium sulfate abdominal constriction test, 125 mg kg^{-1} $MgSO_4$ was injected intraperitoneally in a volume of 10 ml kg^{-1}. Mice were returned to their cubicles and filmed for 20 min after injection. For the spared nerve injury (SNI), surgery (with the sural nerve left intact) was performed under general anesthesia as previously described[23] for the mouse, an adaptation of the rat model[24]. Mice were returned to their home cages until testing. On test days (1, 7 and 14 d postoperative), mice were habituated for 30 min before being filmed for 1 h. For the tail-clip test (TC), a binder clip applying 700 g of force was applied ~1 cm from the base of the tail. The endpoint was a purposeful attempt to remove the clip (head movement toward the tail), usually occurring within 2–3 s (data not shown). Because this assay is sensitive to repeated testing, only one trial was conducted per subject. For the tail-flick test (TF), a radiant heat source was applied to the ventral surface of the tail (~3 cm from the base). The endpoint was reflexive withdrawal of the tail from the stimulus. The intensity of the commercial device (IITC model 33) was set to 5% of maximum output, resulting in mean latencies of approximately 4 s (data not shown). For the tail-withdrawal test (TW), the distal half of the tail was immersed in a temperature-controlled hot water bath (45 °C or 49 °C). The endpoint was reflexive withdrawal of the tail from the water, which occurred in approximately 10 and 5 s, respectively (data not shown). Finally, for the zymosan inflammation test (ZYM), 20 µl of 0.1–10 mg ml^{-1} zymosan was injected subcutaneously into the plantar surface of the right hindpaw, or intra-articularly into the right ankle (n = 6–10 mice per dose at each site). The tibiotarsal joint is marked by two bony prominences that can be felt through the skin (representing the head of the tibial bone and the head of the tarsal bone). Under brief isoflurane-oxygen anesthesia, a 30.5-gauge needle was inserted at a right angle into the joint space. Accuracy of the injections was determined using blue dye in pilot studies. Mice remained in their home cages for 2 h and were then returned to their observation cubicles for 60 min before being filmed for 1 h.

Frame grabbing. For assays involving abrupt withdrawal responses (TC, TF and TW), we grabbed frames during the exhibition of the response itself (pain) and compared these to frames grabbed after the onset of the noxious stimulus but several seconds before the response (no pain). It is arguable whether a more appropriate baseline might be frames taken before the application of the noxious stimulus, but this is mooted by the mostly negative data obtained (**Fig. 2a**). For assays involving longer-lasting nocifensive responses (for example, AA, AITC, CAP, F_{early}, F_{late} and $MgSO_4$), we grabbed frames during the exhibition of such behavior (for abdominal constrictions) or immediately preceding the behavior (for licking, as the face was generally obscured during this behavior; pain); we compared these frames to those grabbed from baseline video (before injection; no pain). For the incision and LAP assays, we grabbed frames randomly from video recorded 1–2 h after surgery (pain) and compared these to frames grabbed 1 d before surgery (no pain). For the CYP and ZYM assays, we grabbed frames randomly 3–4 h after injection (pain) and compared these to frames grabbed from baseline video (before injection; no pain). For the CCI and SNI neuropathic assays, we grabbed frames randomly from video recorded 1, 7 and 14 d after surgery (pain) and compared these to frames grabbed from baseline video (collected one week before surgery; no pain).

For all nociceptive assays described above, pain and no-pain frames were grabbed from digital video (WMV) files, saved as JPG files, and cropped and edited as described above. Edited JPG files were then copied into PowerPoint, one JPG image per slide. A macro obtained on the web (http://www.tushar-mehta.com/powerpoint/randomslideshow/index.htm) was then used to randomize the slide order. Identifications were removed to ensure that subsequent coding was performed in a blinded fashion.

MGS coding. Randomized and unlabeled photos were presented on a large, high-resolution computer monitor, one at a time. For each photo, the scorer assigned a value of 0, 1 or 2 for each of the five AUs: orbital tightening, nose bulge, check bulge, ear position and whisker change. In every case, a score of 0 indicated high confidence of the scorer that the AU was not present. A score of 1 indicated either high confidence of a moderate appearance of the AU or equivocation over whether the AU was present or not. A score of '2' indicated high-confidence detection of a severe appearance of the AU. In cases in which the coder judged that the anatomical feature was obscured, no score was given for that AU. If more than two AUs were obscured, the overall score was not computed for that photo.

In every assay, an average no-pain and an average pain MGS score were calculated, and the existence of a pain face inferred in most cases by a statistically significant (compared to zero) pain MGS minus no-pain MGS mean difference score.

Chemical lesions. Mice were anesthetized with isoflurane-oxygen, shaved and placed in a stereotaxic apparatus such that the skull was horizontal. After incision, drilling through the skull and removal of the dura using a dental needle, 0.1 µl of 10 µg $µl^{-1}$ ibotenic acid (α-amino-3-hydroxy-5-isoazoleacetic acid; Sigma) in PBS or PBS alone (n = 6 mice per condition per site) was injected bilaterally over 1 min using sterilized 0.5-µl Hamilton microsyringes (90-µm tip diameter). The needles were left in place for 5 min to allow diffusion. Wound clips were applied to the scalp to close incisions; these were removed one week after surgery. Mice recovered from anesthesia on a warming plate and were then returned to their home cages. Behavioral testing occurred no less than one week after surgery.

In separate experiments, the anterior cingulate gyrus (ACC), the basolateral amygdala (BLA) and the rostral anterior insula (rAI) were targeted. Coordinates were chosen using the mouse stereotaxic atlas[25] as follows (relative to bregma): for ACC, anterior-posterior (AP) 0.50 mm, dorsal-ventral (DV) −1.75 mm and lateral (L) ±0.25 mm; for BLA, AP −1.46 mm, DV −4.75 mm, L ±2.75 mm; and for rAI, AP 1.80 mm, DV −2.50 mm and L ±2.50 mm.

Behavioral testing occurred on the abdominal constriction ('writhing') test. Mice were videotaped for 30 min after an intraperitoneal injection of 10 ml kg^{-1} of 0.9% acetic acid. The videotape was scored on the MGS and also for abdominal constrictions using a sampling strategy[14], in which the presence or absence of abdominal constrictions or 'writhes' (characteristic lengthwise stretches of the torso with a concomitant concave arching of the back) in a 5-s period at 20-s intervals (90 total samples) was scored

© 2010 Nature America, Inc. All rights reserved.

by a observer blinded to pain status. After behavioral testing, mice were anesthetized with pentobarbital (100 mg kg^{-1}) and perfused with 4% paraformaldehyde in 0.1 M phosphate buffer and sodium nitrite (0.1%). Brains were removed, processed in 4% paraformaldehyde and then 30% sucrose, cryosectioned into 25-µm-thick slices and stained with cresyl violet to confirm targeting. Based on this, two of six rAI lesions were judged to have missed the rAI, and data for these mice were analyzed as a separate group.

Morphine experiment. Morphine analgesia was assessed in the CYP model using the highest dose of CYP, 400 mg kg^{-1}. CYP was administered intraperitoneally in a volume of 10 ml kg^{-1} after baseline filming. Saline or morphine sulfate (5 or 10 mg kg^{-1}) was administered subcutaneously 3.25 h after CYP injection (n = 4–6 mice per dose), allowing 15 min before the start of a 30-min post-drug injection videotaping. One group of mice was injected with 10 mg kg^{-1} morphine only.

Acetaminophen experiment. Separate groups of naive mice (n = 4–8 mice per drug per assay) were tested on the MGS and for mechanical allodynia after injection of 100 µg 5 mg ml^{-1} zymosan. Mechanical sensitivity was measured in the von Frey test using the up-down staircase method[26]. Mice were placed on a metal mesh floor in small plexiglas cubicles (9 × 5 × 5 cm high), and eight calibrated von Frey fibers (0.007–1.4 g) were applied to the plantar surface of the hindpaw until they bowed. The threshold force required to elicit withdrawal of the paw (median 50% paw withdrawal) was determined. Data presented are from both left and right hind paws averaged together as no laterality effects were noted.

For both measures, baseline data were collected in a 30-min period before zymosan administration. Post-zymosan (but pre-acetaminophen) data were collected in both cases in a 30-min period starting 4 h after zymosan administration. Immediately thereafter, mice were injected subcutaneously with either 300 mg kg^{-1} acetaminophen or vehicle (polyethylene glycol; 10 ml kg^{-1}). This dose was chosen as it is the recommended mouse postoperative analgesic dose at our institution (http://www.mcgill.ca/files/researchoffice/101_02-Rodent_Analgesia.doc; 8 July 2009), but we note that convincing demonstrations of efficacy of acetaminophen in inflammatory assays other than the (modest intensity) acetic acid abdominal constriction test are rare[27,28]. Thirty min later, a 30-min-long videotaping or von Frey testing occurred again, to assess the possible analgesic or antiallodynic affects of acetaminophen. Percentage analgesia scores of acetaminophen-treated mice presented in **Figure 3b** were calculated with respect to the maximum possible effect for each mouse, as follows: percentage analgesia = ((post-zymosan threshold – post-acetaminophen threshold) / (baseline threshold – post-zymosan threshold)) × 100. Full results including vehicle-group data are available in **Supplementary Figure 3**.

'Migraine mouse' experiment. The MGS was applied to a transgenic mouse model of familial hemiplegic migraine, type 1 (FHM-1). As described previously[13], using a gene targeting approach and applying Cre-*lox* technology to remove the neomycin-resistance cassette, a knockin mouse was developed carrying the *Cacna1a* missense mutation (encoding a protein with the R192Q mutation) that in humans causes FHM-1 (ref. 29). Gain-of-function effects have been observed in this mutant, including an increase in neuronal $Ca_v2.1$ current density, enhanced cortical neurotransmission and enhanced propensity for cortical spreading depression[13,30], the likely mechanism underlying the migraine aura[31]. One can speculate, therefore, that the mice carrying the *Cacna1a* missense mutation (encoding a protein with the R192Q mutation) might actually experience migraine headaches[32]. We have undertaken a systematic characterization of the behavioral phenotype of these mutants (unpublished data; M.L. Chanda *et al.*).

Naive, adult (7–21-week-old) *Cacna1a* knockin mice and their wild-type counterparts were placed in transparent plexiglas cylinders (3-cm diameter; 10-cm long) in front of video cameras as described above. After a 30-min habituation, mice were briefly removed and injected intraperitoneally with either 50 mg kg^{-1} of rizatriptan or saline (n = 4–6 mice per drug per genotype). Frames were captured from video at random intervals over 2 h, yielding approximately 45 photos per mouse. In this case, there can be no 'baseline' for comparison; we believe that the cylinder restraint itself may have triggered migraines in a subset of mutant mice, leading to increased MGS scores overall in this genotype. Accompanying the pain face in mutant mice are other head-specific behaviors that are considerably more frequent in mutant mice and normalized by rizatriptan at this same dose (unpublished data; M.L. Chanda *et al.*).

Statistical analysis. All statistical analyses were performed using Systat v.11 (SPSS Inc.), with a criterion $\alpha = 0.05$, except for the ICC in **Supplementary Figure 1b**, which was calculated using SPSS v. 17. In most cases where MGS mean difference scores were calculated, those scores were compared to zero using a one-sample, one-tailed Student's t-test. Correction for multiple comparisons in **Figure 2a** was performed using the false discovery rate technique[33], although we note that experiments were run separately, not concurrently. Dose- or stimulus intensity–dependency (**Figs. 2b** and **3a**) was analyzed by one-way ANOVA, followed where appropriate by Dunnett's case-comparison posthoc test (comparing to vehicle). Group differences in the lesion experiments (**Fig. 2c**) were assessed by one-way ANOVA followed by Tukey's posthoc test. Effects of genotype and drug in the 'migraine mouse' experiment (**Fig. 3c**) were analyzed by two-way ANOVA, followed by Student's t-tests.

16. Shrout, P.E. & Fleiss, J.L. *Psychol. Bull.* **86**, 420–428 (1979).
17. Mogil, J.S. *et al. Pain* **80**, 67–82 (1999).
18. Mogil, J.S. *et al. Pain* **126**, 24–34 (2006).
19. Bennett, G.J. & Xie, Y.-K. *Pain* **33**, 87–107 (1988).
20. Bon, K., Lichtensteiger, C.A., Wilson, S.G. & Mogil, J.S. *J. Urol.* **170**, 1008–1012 (2003).
21. Pogatzki, E.M. & Raja, S.N. *Anesthesiology* **99**, 1023–1027 (2003).
22. Brennan, T.J., Vandermeulen, E.P. & Gebhart, G.F. *Pain* **64**, 493–501 (1996).
23. Shields, S.D., Eckert, W.A. III & Basbaum, A.I. *J. Pain* **4**, 465–470 (2003).
24. Decosterd, I. & Woolf, C.J. *Pain* **87**, 149–158 (2000).
25. Franklin, K.B. & Paxinos, A. *The Mouse Brain in Stereotaxic Coordinates*, 2nd ed. (Academic Press, 2003).
26. Chaplan, S.R., Bach, F.W., Pogrel, J.W., Chung, J.M. & Yaksh, T.L. *J. Neurosci. Methods* **53**, 55–63 (1994).
27. Dickinson, A.L., Leach, M.C. & Flecknell, P.A. *Lab. Anim.* **43**, 357–361 (2009).
28. Wilson, S.G. *et al. J. Pharmacol. Exp. Ther.* **305**, 755–764 (2003).
29. Ophoff, R.A. *et al. Cell* **87**, 543–552 (1996).
30. Tottene, A. *et al. Neuron* **61**, 762–773 (2009).
31. Goadsby, P.J., Lipton, R.B. & Ferrari, M.D. *N. Engl. J. Med.* **346**, 257–270 (2002).
32. Xue, M. & Rosenmund, C. *Neuron* **61**, 653–654 (2009).
33. Benjamini, Y. & Hochberg, Y. *J. Roy. Statist. Soc. Ser. B.* **57**, 289–300 (1995).

© 2010 Nature America, Inc. All rights reserved.

doi:10.1038/nmeth.1455