Data Descriptor: Five years of phenology observations from a mixed-grass prairie exposed to warming and elevated CO₂

Melissa Reyes-Fox¹⁺, Heidi Steltzer²⁺, Daniel R. LeCain³ & Gregory S. McMaster⁴

Atmospheric CO₂ concentrations have been steadily increasing since the Industrial Era and contribute to concurrent increases in global temperatures. Many observational studies suggest climate warming alone contributes to a longer growing season. To determine the relative effect of warming on plant phenology, we investigated the individual and joint effects of warming and CO₂ enrichment on a mixed-grass prairie plant community by following the development of six common grassland species and recording four major life history events. Our data support that, in a semi-arid system, while warming advances leaf emergence and flowering, it also expedites seed maturation and senescence at the species level. However, the additive effect can be an overall lengthening of the growing and reproductive seasons since CO₂ enrichment, particularly when combined with warming, contributed to a longer growing season by delaying plant maturation and senescence. Fostering synthesis across multiple phenology datasets and identifying key factors affecting plant phenology will be vital for understanding regional plant community responses to climate change.

Design Type(s) | time series design  ●  parallel group design
Measurement Type(s) | life cycle stage  ●  climate data  ●  volumetric soil water
Technology Type(s) | phenological observation  ●  meteorological station  ●  water content profile probe
Factor Type(s) | atmospheric carbon dioxide  ●  air temperature
Sample Characteristic(s) | State of Wyoming  ●  prairie  ●  Artemisia frigida  ●  Bouteloua gracilis  ●  Hesperostipa comata  ●  Koeleria macrantha  ●  Pascopyrum smithii  ●  Sphaeralcea coccinea

¹USDA-ARS, Soil Plant Nutrient Research Unit and Northern Plains Area, Fort Collins, Colorado 80526, USA.
²Department of Biology, Fort Lewis College, Durango, Colorado 81301, USA.
³USDA-ARS, Rangeland Resources Research Unit, Fort Collins, Colorado 80526, USA.
⁴USDA-ARS, Agricultural Systems Research Unit and Northern Plains Area, Fort Collins, Colorado 80526, USA. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to M.R.-F. (email: melissa.reyes@ars.usda.gov) or to H.S. (email: steltzer_h@fortlewis.edu).
Background & Summary
Climatic change factors, including increasing temperatures and rising CO₂ levels, have been shown to affect plant phenology1–3. We investigated the effects of warming and CO₂ enrichment on the timing and duration of four key developmental phases (phenophases) and found, in a mixed-grass prairie, the indirect effects of warming and CO₂ enrichment on soil water availability play a central role in governing whether plants expedite or prolong development. If sufficient water resources exist, and temperatures remain above minimum thresholds, species will persist later into the season. If not, they often mature and senesce more quickly. Independent shifts in the magnitude and direction of life histories are species-specific, which in our system means the duration of the growing and reproductive seasons can be attributed to the response of one or two species per year (not necessarily the same species), and not a community-wide tendency to either extend or shorten growth and reproduction. Shifts in species’ life histories can influence plant community composition by disrupting timing sensitive relationships between plants and their associated pollinators and herbivores, and by decreasing the level of species’ complementarity. Variations in the timing of resource use can have detrimental effects on plant communities by facilitating invasive species colonization and limiting nutrient uptake/cycling via reductions in terrestrial biomass production4–6.

Warming has the potential to extend the growing and reproductive seasons by promoting earlier leaf and flower emergence2–7. However, those species showing significant advancements in leaf emergence and flower production in response to warming also matured and senesced earlier. We provide evidence that warming causes earlier leaf emergence but the magnitude of this advancement was not as great as the delay in senescence that resulted from exposure to combined warming and CO₂ enrichment8 (see Reyes-Fox and Steltzer et al. 2014 for the full Nature article). Mean growing season duration was lengthened by 6.2 (+/− 8.0) days under warming alone versus 14.2 (+/− 7.0) days under combined warming and CO₂ enrichment, indicating warming may not be the only, or even primary, phenological cue in prolonging the growing season8,10. CO₂ enrichment has been shown to play an important role in limiting stomatal aperture and thus minimizing water loss via transpiration. Experimental research shows evidence that the water savings resulting from CO₂ enrichment counteracts the desiccating effect of moderate warming8,11. Where direct effects of CO₂ enrichment on plant phenology are concerned, the conclusions vary greatly. A number of studies focus on measures of reproductive investment such as floral abundance and mass, seed viability, etc. (refs 12–14), but not on the effect on the timing of leaf production or senescence. Research suggests elevated CO₂ affects the timing of plant reproduction variably, expediting flowering in some species and delaying them in others15 and it has been shown to stimulate growth directly by enhancing photosynthetic rates and thus biomass production16–18. Many plants must reach a minimum size before a secondary cue (e.g., daylength) will trigger flowering. CO₂ can directly affect flowering time via this mechanism but the magnitude and direction of this effect is difficult to predict19,20.

The authors present this phenology dataset and the accompanying microclimate data, which have reuse value in the fields of plant ecology and climate change. The raw data have implications for developing models that use environmental cues to predict the timing of phenology events for important developmental stages to test whether common patterns emerge from multiple datasets across different sites. There is also potential to use the data to conduct sensitivity analyses to determine thresholds for volumetric soil water content (VSWC) in different systems or to create thermal time estimates of species growing degree-day (GDD) responses for different developmental stages. Synthesis analyses across multiple phenology datasets will be useful in identifying key factors affecting plant phenology (e.g., late season VSWC in our case) and illuminating areas where uncertainty exists. The data will also be relevant to land managers interested in the timing of annual life cycle events to determine best management practices (prescribed burning, for example) and ranchers who rely on predictions of aboveground biomass production to determine stocking rates. Phenology data can also be used by those afflicted by seasonal allergies, or who are interested mutualisms, designations of migratory corridors, or any other area where the timing of life history events is important.

Methods
Site description
The Prairie Heating and Carbon Dioxide Enrichment (PHACE) experiment, initiated in 2006, is located west of Cheyenne, WY, USA at the USDA-ARS High Plains Grasslands Research Station in the U.S. Great Plains (41° 11’ N, 104° 54’ W, elevation 1,930 m). This is a Northern mixed-grass prairie ecosystem, with a plant community comprised of 55% cool-season grasses, 25% warm-season grasses, and 20% sedges, forbs, and small shrubs. Total annual precipitation averages 38.5 cm and mean daily air temperatures range from −2.5 °C in January to 17.5 °C in July. The average wind speed is 6 m s⁻¹ with gusts up to 35 m s⁻¹. The site is comprised of two distinct soil types: an Ascalon Variant Loam (fine-loamy, mixed mesic) at the north end of the field and an Altvan Loam (fine-loamy, mixed mesic) at the south end. The 2.4 hectare site has a history of moderate grazing from 1928 until 2005, when fences were installed to prevent cattle from entering.
Experimental design
The experiment includes two levels of temperature (ambient and warmed, 1.5/3.0 °C warmer day/night, treatments t and T, respectively) and two levels of atmospheric CO₂ concentrations (ambient 385 ppmv and elevated 600 ppmv CO₂, treatments c and C, respectively) in a factorial combination with five replicate plots per treatment (ct, cT, Ct, and CT) for a total of 20 plots. Differential day/night time temperatures were implemented because minimum temperatures are predicted to increase at a faster rate than maximum temperatures\textsuperscript{21}. Treatment combinations of ambient and elevated temperatures and CO₂ concentrations are modeled after current climate change projections based on moderate, continued increases in CO₂ for the end of this century\textsuperscript{22}. Warming and elevated CO₂ treatments were randomly assigned to the 3.3 m diameter circular plots (Fig. 1 depicts an outline of the approach used in this experiment). T-FACE technology for increasing temperature was implemented in spring of 2007 and warmed plots year round for the duration of the experiment\textsuperscript{23}. Dummy heaters were installed in non-heated plots to eliminate response differences that may result from shading or other influences caused by the heating apparatuses. Free air CO₂ enrichment (FACE) technology was used for enriching CO₂ and began in 2006 (ref. 24). The CO₂ fumigation system ran continuously during the daylight hours of the growing season (typically from about 1 April to 15 November), but was turned off after the plants ceased growing each year. Figure 2 depicts the experimental layout of the entire PHACE experiment; including the 10 irrigated plots (where phenology data was collected but is not reported here).

Observations
The timing of four life cycle events that determine the start and end of species’ active and reproductive periods (leaf emergence, flower production, seed maturation, and canopy senescence) was observed weekly for six common species. The most abundant species in each growth form were chosen, including the one sub-shrub Artemisia frigida, L. (ARFR); a warm-season grass Bouteloua gracilis, Lag. ex Griffiths (BOGR); three cool-season grasses Hesperostipa comata, (Elías) Barkworth (HECO), Koeleria macrantha, (Ledeck.) Schult (KOMA), and Pascopyrum smithii, (Ryd.) Á. Löve (PASM); and a widespread forb Sphaeralcea coccinea, (Nutt.) Rydb (SPCO). Leaf emergence was characterized by the first new, green leaf to appear on a shoot. Flower production was reached when the first open flower (ARFR and SPCO) or inflorescence emerged from leaf sheath (BOGR, HECO, PASM, KOMA). Maturation was characterized as spikes subtending below the inflorescence which detach easily for BOGR; no green colour on seed heads for PASM and KOMA; wilted/dead flowers for ARFR and SPCO; seeds are unsheathed in HECO and are easily separated from the peduncle/stalk of the inflorescence. Senescence was reached when less than 10% green colour in the entire plant remained and plant growth ceased. Some species (i.e. ARFR) can overwinter and therefore will never be entirely brown.

To conduct phenology observations, a sample grid was placed over the centre area of the plot, and an access platform was used so the observer could hover over each of the 20 plots. The 3.3 m diameter circular plots were subdivided into two halves: one supporting the native prairie vegetation and the other dedicated to a separate invasive plant study. The far end of the wooden platform was set upon a metal flange located between the northern mixed-grass prairie study area and the weed observation area. The other end of the platform fell outside the CO₂ injection tubing which circled the perimeter of the plot, thereby avoiding damage to the plots. Positioning the platform above the plots ensured the observer could view the plants up close. During years 2007–2009, the base of each individual was marked with a headpin and those same marked individuals were re-visited every week and their developmental phase was recorded for the duration of the growing season. Non-grass species (i.e., ARFR and SPCO) were not marked because they were easily identified. To ensure spatial variability we tried to identify one individual per species in each of the 24 quadrats of the sample grid. New individuals were numbered sequentially and added to the database. Typically at least one individual from BOGR, PASM, and HECO were found in each of the quadrats. Sample sizes for KOMA, ARFR and SPCO varied greatly between plots with an average of five individuals per plot per species. From 2007–2009, if a particular individual plant did not re-emerge after over-wintering it was replaced with another individual plant selected from nearby. If an alternate plant could not be located, the individual was removed from the list. Using maps for each plot, ID numbers for new individuals were assigned and arranged on the maps and given a distinguishing characteristic (i.e., bold or italics) depending on their current phenophases. This protocol was also followed if previously marked individuals were missing or senesced. In 2010 and 2011, the methods changed slightly and the plot maps were no longer used to track phenostages. Instead of tracking phenostages of individual plants for each target species, we used a ‘threshold’ to identify when a species

Figure 1. Experimental workflow diagram of PHACE phenology study.
had reached a particular phenostage. For instance, we surveyed the sample area and when we could see 10 BOGR individuals that had reached leaf emergence, we recorded it. Less common species like ARFR had a threshold of 5. In 2010 and 2011 the same phenostage criteria were used except heading in grasses was also defined, where possible. Over the course of the study, observers differentiated between senesced and non-existent individuals.

**Data Records**

This was an experimental study where observations were conducted from 2007 through 2011, and where no samples were collected from within the plots nor were laboratory analysis conducted. Data outputs for plant individuals, climate and VSWC were generated and the reader is referred to corresponding data files (Table 1) where primary data and metadata are archived. All data were entered into an Excel
Table 1. Location and description of primary databases.

| Data Record | Source | Sample | Number of data values | Temporal range | Protocol 1 | Protocol 2 | Data file name | Repository | Size |
|-------------|--------|--------|-----------------------|----------------|------------|------------|----------------|------------|------|
| 1           | Reyes-Fox, M., Steltzer, H., LeCain, D. R., McMaster, G. S. | Phenology | 117635 | 2007–2011 | Plot-level phenology observations | Method to identify treatment differences | PHACEphenology_database_final_forSD.xlsx | Dryad | 587.9 Kb |
| 2           | Reyes-Fox, M., Steltzer, H., LeCain, D. R., McMaster, G. S. | Climate | 359630 | 2007–2011 | Site-level climate data | Method to test interannual variability | PHACEclimate_database_final_forSD.xlsx | Dryad | 2.299 Mb |
| 3           | Reyes-Fox, M., Steltzer, H., LeCain, D. R., McMaster, G. S. | Volumetric Soil Water Content | 7307 | 2007–2011 | Plot-level soil water measurements | Method to identify treatment differences | PHACEswc_database_final_forSD.xlsx | Dryad | 76.64 Kb |

Phenology
From 2007–2011 all plots were clearly delimited and weekly phenology observations were recorded within each of the 20 plots from mid-March to early November. We used observational data from this time period to determine changes in growing and reproductive season length. Data are presented annually and across years for the duration of species’ active and reproductive periods (Data Citation 1; PHACEphenology_database_final_forSD.xlsx). The start and end of the growing season were characterized by the mean across replicate plots for leaf emergence by the first species to leaf and for canopy senescence by the last species, respectively. Similarly, the start and end of the reproductive season were characterized by the date when the first species flowered and when the last species reached seed maturation, respectively. Table 2 (available online only), reports replicate means in DOY for the 4 observed phenostages by year and by species. Standard errors and sample sizes are also provided.

Climate and VSWC
Mean daily temperature and precipitation were calculated based on half hourly data from a meteorological station (HOBO, Onset, Inc., MA) at the field site (Data Citation 1; PHACEclimate_database_final_forSD.xlsx). In some cases (namely in 2011) missing data was replaced by data from a proximate met station that recorded hourly data instead of half hourly data. The substitute met station was the High Plains Grassland Research Station located approximately 1 km south of the PHACE pasture. It was a Campbell Scientific system and therefore had different sensors and datalogger programming. In the winter of 2009–2011, precipitation amounts were adjusted to account for moisture received as snowfall but this correction was done to daily averages, not the half hourly data. In 2009–2011, daily precipitation totals were adjusted to account for differences in snowfall versus rainfall using a snowfall adapter at the HPGRS. Prior to 2009, an accurate means of measuring snow moisture equivalents was not available. In our case, site-specific climate data was used to depict seasonal variation in precipitation and air temperature among years. When placed in a historical context of the last century, below average precipitation fell in 2007, 2008, and 2010 with above average precipitation falling in 2009 and 2011. 2010 was the warmest and driest of the five years and 2011 was the wettest year, however temperatures in 2009 were the coolest.

In each plot, the VSWC was measured hourly at 10, 20, 40, 60 and 80 cm depths (EnviroSMART probe; Sentek Sensor Technologies, Stepney, Australia). However, we present only the fall soil water measurements here, since significant treatments differences at this period of the growth season were determined to affect plant phenology. Daily means were calculated for soil water content at the primary rooting depth (5–25 cm) by averaging the values for the sensors at 10 and 20 cm depth (Data Citation 1; PHACEswc_database_final_forSD.xlsx). Interannual variation in microclimate interacted with species phenology to produce dramatic differences in VSWC, particularly at the end of the season, among years.

Technical Validation
Quality assurance and control
From 2007–2009 individual plants were marked, species were color-coded and these marked individuals were re-visited to avoid misidentification. This was particularly helpful when the plants were first emerging and looked similar. If the pins were non-detectable, the observer referred to the field maps, where the location of each individual plant was recorded. If the specific location of a particular plant was not found, an alternate, nearby individual was marked. We minimized misidentification by executing thorough training in species identification. This included a combination of in-situ plant identification

spreadsheet and Proc Mixed (SAS version 9.2, 2008, SAS Institute, Cary, NC) was used to analyse the data (see Reyes-Fox and Steltzer et al. 2014 for full disclosure of statistical analyses). Plot markers and infrastructure were removed at the conclusion of the experiment.
training by data observers and use of photographic plant guides for visuals of phenological stages for each species (Fig. 3 gives an example for two species). To ensure quality of data collection (i.e., observations), we also harvested sample specimens of each species from outside of the plots, archiving them in their vegetative and reproductive states to reference as needed. Samples were not collected for physical or chemical analysis of any kind. In 2010 and 2011, instead of following the same individual plants, the timing of an event was achieved when a minimum number of individuals for each species within a plot had completed a life history event, representing the median value. Despite this difference in approaches, both methods represented a central tendency to quantify event timing across multiple individuals per plot. We performed a direct one to one comparison of treatment means presented in Table 2 (available online only) for Method 1 (2007–2009) versus Method 2 (2010–2011) at each growth stage averaged across species for each method. A linear regression analysis plotting one method on the x-axis and the other on the y-axis against produced a $R^2$ value of .98.

Data were recorded and updated weekly, and were checked for inconsistencies. We also made certain data fields were accurately labelled. To ensure treatments (i.e., heaters and CO$_2$ injection system) were operating properly, data loggers were installed at PHACE and were routinely monitored$^{11,25}$. Species-level phenological trends varied from year to year but the likelihood that this variation was due to observer error is low since a standard reporting protocol was implemented and data were collected over the course of the five year study by the same 3 persons. Also, treatment level effects were not as variable as the within-plot species-level variation.

Figure 3. Pictorial index of plant phenophases for (a) C-3 grass (*Pascopyrum smithii*) and (b) C-4 grass (*Bouteloua gracilis*). Photo credits Daniel LeCain and Julie Kray.
T-FACE technology for increasing temperature began on April 10th 2007, after leaf emergence by the cool-season grasses and shrubs. Since warming began after the first species leafed, leaf emergence data were omitted for all species in 2007 and growing season length was not calculated. Similarly, some species did not flower in all years. In 2009, flower production measurements were not collected from the control (ct) and CO₂ enriched plots (Ct) for BOGR, and seed maturation was not recorded for ARFR or SPCO in any of the treatments. Therefore reproducible season length could not be determined for this particular year.

References
1. CaraDonna, P. J., Iler, A. M. & Inouye, D. W. Shifts in flowering phenology reshape a subalpine plant community. *Proceedings of the National Academy of Science* 11, 4916–4921 (2014).
2. Tylianakis, J. M., Didham, R. K., Bascompte, J. & Wardle, D. A. Global change and species interaction in terrestrial ecosystems. *Ecology Letters* 11, 1351–1363 (2008).
3. Parmesan, C. Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology and Evolutionary Systems* 37, 637–669 (2006).
4. Steltzer, H. & Post, E. S. Seasons and life cycles. *Science* 324, 886–887 (2009).
5. Cleland, E. E., Chiariello, N. R., Loarie, S. R., Mooney, H. A. & Field, C. B. Diverse responses of phenology to global changes in a grassland ecosystem. *Proceedings of the National Academy of Science* 103, 13740–13744 (2006).
6. Post, E. S., Pedersen, C., Wilmers, C. C. & Forchammer, M. C. Phenological sequences reveal aggregate life history response to climate warming. *Ecology* 89, 363–370 (2008).
7. Hollister, R. D., Webber, P. J. & Bay, C. Plant response to temperature in northern Alaska: implications for predicting vegetation change. *Ecology* 86, 1562–1570 (2005).
8. Reyes-Fox, M. et al. Elevated CO₂ further lengthens growing season under warming conditions. *Nature* 510, 259–262 (2014).
9. Sherry, R. A. et al. Divergence of reproductive phenology under climate warming. *Proceedings of the National Academy of Science* 104, 198–202 (2007).
10. Cleland, E. E., Chuine, I., Menzel, A., Mooney, H. A. & Schwartz, M. D. Shifting plant phenology in response to global change. *Trends in Ecology and Evolution* 22, 357–365 (2007).
11. Morgan, J. A. et al. CO₂ grasses prosper as carbon dioxide eliminates desiccation in warmed semi-arid grassland. *Nature* 476, 202–205 (2011).
12. Jabonski, L. M., Wang, X. & Curtis, P. S. Plant reproduction under elevated CO₂ conditions: a meta-analysis of reports on 79 crop and wild species. *New Phytologist* 156, 9–26 (2002).
13. Lake, J. C. & Hughes, L. Nectar production and floral characteristics of *Tropaeolum majus* L. growth in ambient and elevated carbon dioxide. *Annals of Botany* 84, 535–541 (1999).
14. Edwards, G. E., Clark, H. & Newton, P. C. D. The effects of elevated CO₂ on seed production and seedling recruitment in a sheep grazed pasture. *Oecologia* 127, 383–394 (2001).
15. Springer, C. J. & Ward, J. K. Flowering time and elevated atmospheric CO₂. *New Phytologist* 176, 243–255 (2007).
16. Isao, S. B., Allen, S. G. & Kimball, B. A. Growth response of water lily to atmospheric CO₂ enrichment. *Aquatic Botany* 37, 87–92 (1990).
17. Niu, G., Heins, R. D., Cameron, A. C. & Carlson, W. H. Day and night temperatures, daily light integral and CO₂ enrichment affect growth and flower development of pansy. *Journal of the American Society of Horticultural Science* 125, 436–441 (2000).
18. Zimmerman, R. C., Kohrs, D. G., Steller, D. L. & Alberte, R. S. Impacts of CO₂ enrichment on productivity and light requirements of eelgrass. *Plant Physiology* 115, 599–607 (1997).
19. Johnston, A. & Reekie, E. Regardless of whether rising atmospheric carbon dioxide levels increase air temperature, flowering phenology will be affected. *International Journal of Plant Science* 169, 1210–1218 (2008).
20. Rathcke, B. & Lacey, E. P. Phenological patterns of terrestrial plants. *Annual Review of Ecological Systems* 16, 179–214 (1985).
21. Trenberth, K. E. et al. in Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (eds Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Areyt, M. Tignor & H.L. Miller.) (Cambridge University Press, Cambridge, United Kingdom & New York, NY, USA, 2006).
22. Farquhar, G. D. et al. IPCC. Working Group 1: The Scientific Basis, Chapter 3. The carbon cycle and atmospheric carbon dioxide, 183–287 (2001).
23. Kimball, B. A. et al. Infrared heater arrays for warming ecosystem field plots. *Global Change Biology* 14, 309–320 (2008).
24. Miglietta, F. et al. Free-air CO₂ enrichment (FACE) of a poplar plantation: the POPFACE fumigation system. *New Phytologist* 150, 465–476 (2001).
25. LeCain, D. et al. Microclimatic Performance of a Free-Air Warming and CO₂ Enrichment Experiment in Windy Wyoming, USA. *PLoS ONE* 10, e0116834 (2015).

Data Citation
1. Reyes-Fox, M. et al. Dryad Digital Repository http://dx.doi.org/10.5061/dryad.267d2 (2016).

Acknowledgements
The following individuals contributed to the design, installation and maintenance of the Prairie Heating and CO₂ Enrichment (PHACE) project: J. Morgan, D. Blumenthal, D. Smith, E. Pendall, D. Williams and E. Hardy. L. Griffith, A. Hansen, K. Corp, V. Banuelos, G. Tinnin, C. Brooks, and M. Nix assisted with the plant identification and collection of phenology data. M. West, M. J. Trlica, A. Andales, and D. Milchunas provided technical advice. Funding for this work was supported by the US Department of Agriculture Agricultural Research Service (USDA-ARS) Climate Change, Soils & Emissions Program, by the US Department of Energy’s Office of Science through the Terrestrial Ecosystem Science Program, by the National Science Foundation (DEB no. 1021559) and by Colorado State University.

Author Contributions
M.R.-F. and G.S.M. designed the research. M.R.-F. and D.R.L. conducted the observations. D.R.L. managed the day-to-day operations for the PHACE experiment. H.S. and M.R.-F. analysed the data and wrote the manuscript. M.R.-F had full access to all the data in the study and takes responsibility for the
integrity of the data. H.S. assumes responsibility for the accuracy of the data analysis. All authors contributed to revision of the manuscript.

Additional Information
Table 2 is only available in the online version of this paper.

Competing financial interests: The authors declare no competing financial interests.

How to cite: Reyes-Fox, M. et al. Five years of phenology observations from a mixed-grass prairie exposed to warming and elevated CO₂. Sci. Data 3:160088 doi: 10.1038/sdata.2016.88 (2016).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0

Metadata associated with this Data Descriptor is available at http://www.nature.com/sdata/ and is released under the CC0 waiver to maximize reuse.

© The Author(s) 2016