Differentiation of primate primordial germ cell-like cells following transplantation into the adult gonadal niche

Enrique Sosa1, Di Chen1, Ernesto J. Rojas1, Jon D. Hennebold2,3, Karen A. Peters4, Zhuang Wu5, Truong N. Lam5, Jennifer M. Mitchell6, Meena Sukhwani4, Ramesh C. Tailor7, Marvin L. Meistrich5, Kyle E. Orwig4, Gunapala Shetty5 & Amander T. Clark1

A major challenge in stem cell differentiation is the availability of bioassays to prove cell types generated in vitro are equivalent to cells in vivo. In the mouse, differentiation of primordial germ cell-like cells (PGCLCs) from pluripotent cells was validated by transplantation, leading to the generation of spermatogenesis and to the birth of offspring. Here we report the use of xenotransplantation (monkey to mouse) and homologous transplantation (monkey to monkey) to validate our in vitro protocol for differentiating male rhesus (r) macaque PGCLCs (rPGCLCs) from induced pluripotent stem cells (riPSCs). Specifically, transplantation of aggregates containing rPGCLCs into mouse and nonhuman primate testicles overcomes a major bottleneck in rPGCLC differentiation. These findings suggest that immature rPGCLCs once transplanted into an adult gonadal niche commit to differentiate towards late rPGCs that initiate epigenetic reprogramming but do not complete the conversion into ENO2-positive spermatogonia.

https://doi.org/10.1038/s41467-018-07740-7
Germline cells are essential for fertility and passing DNA from one generation to the next. In each generation, germ cell development begins around the time of embryo implantation with the differentiation of founding progenitors called primordial germ cells (PGCs). PGCs are transient and in the appropriate environment will subsequently advance in differentiation towards oogonia in females and pro-spermatogonia in males. In an inappropriate environment, however, the latent pluripotency program can be reactivated leading to germ cell tumors including teratomas. Moreover, the abnormal specification of PGCs has the potential to impact the quality of the entire cohort of germ cells in the adult gonad given that after PGC specification no other cell type can contribute to the germline. Therefore, understanding the biology of PGCs has important implications for future reproductive success and child health.

One of the most exciting models for understanding human PGC development is the pluripotent stem cell model and differentiation into PGC-like cells (PGCLCs) in vitro 1–5. Directed differentiation protocols for generating human PGCLCs (hPGCLCs) result in the formation of so-called early PGCs which are equivalent to PGCs at around week 3 of human embryo development. Early PGCs in the primate cynomolgus ( cyno) macaque are triple positive for SOX17, PRDM1, and TFAP2C, while being negative for the late stage PGC markers VASA and DAZL 6. A recent study has demonstrated that female human primate PGCLCs create VASA negative cells equivalent to early-stage PGCs 1,2. In the cynomolgus macaque, VASA protein expression is induced in cynoPGCs at around Carnegie Stage (CS) 12 (Days 26–30) 6. To evaluate this in rhesus macaque embryos, we collected time-mated CS12 rhesus embryos (n = 3) (Fig. 1a). Examination of transverse histological sections in the region of the aorta–gonad–mesonephros (Fig. 1a, box, and arrow), revealed gonadal ridge epithelium, (Fig. 1b, arrow-heads), and not the DA. Therefore, the DA SOX17 positive cells are most likely the hemogenic endothelium 21. This also indicates that TFAP2C is a discriminating marker of primate PGCs in this region of the embryo at this developmental stage 6. Triple IF staining of SOX17, TFAP2C, and the transcription factor PRDM1, all markers of early PGCs, identifies triple positive rPGCs in the dorsal mesentery as well as rPGCs closely associated with the gonadal ridge epithelium (Fig. 1d). Furthermore, all TFAP2C positive rPGCs in CS12 co-express the pluripotency transcription factor OCT4 (Fig. 1e). To determine whether rPGCs at CS12 correspond to early rPGCs (VASA negative) or late rPGCs (VASA positive) 6, we stained for VASA and the rPGC marker TFAP2C. Our data indicate that the majority of CS12 rPGCs were located in either the dorsal mesentery or the genital ridge epithelium, and the majority of TFAP2C positive rPGCs were also VASA positive (Supplementary Figure 1a, b). In addition, rare VASA negative TFAP2C positive cells were also identified (Fig. 1e, white arrows). As expected SOX2 was not expressed in any TFAP2C positive rPGCs (Fig. 1e).

In order to confirm that the testicular xenotransplantation bioassay could be used as an important reporter for germline competency despite the lack of apparent differentiation, Hermann and colleagues 18,19 showed that homologous transplantation of rhesus macaque testicular cells into recipients depleted of spermatogonial stem cells prior to transplantation promotes spermatogenesis from donor cells 16,20. Furthermore, not only were the donor SSCs competent to undergo complete spermatogenesis the donor-derived sperm were competent to fertilize rhesus macaque oocytes and give rise to donor-derived embryos 20. It is unknown how less mature rhesus macaque germ cell types will respond in this assay.

In the current study, we differentiated riPSCs to early (VASA negative) rPGCLCs that we characterize as being similar to bona fide embryonic rPGCs younger than 28 days of embryo development post-fertilization. Following xenotransplantation into irradiated nude mice or homologous transplantation into irradiated rhesus macaques, we show that the seminiferous tubules environment supports the survival of rPGCLCs and promotes further differentiation to a VASA-positive state. Taken together, transplantation to the seminiferous tubule environment promotes rPGCLC differentiation beyond what can currently be achieved in vitro.

**Results**

**Rhesus PGCLCs equate to rPGCs younger than embryonic day 28.** To date, all directed differentiation strategies to generate primate PGCLCs create VASA negative cells equivalent to early-stage PGCs. In the cynomolgus macaque, VASA protein expression is induced in cynoPGCs at around Carnegie Stage (CS) 12 (Days 26–30) 6. To evaluate this in rhesus macaque embryos, we collected time-mated CS12 rhesus embryos (n = 3) (Fig. 1a). Examination of transverse histological sections in the region of the aorta–gonad–mesonephros (Fig. 1a, box, and arrow), revealed gonadal ridge epithelium, (Fig. 1b, arrow-heads), and not the DA. Therefore, the DA SOX17 positive cells are most likely the hemogenic endothelium 21. This also indicates that TFAP2C is a discriminating marker of primate PGCs in this region of the embryo at this developmental stage 6. Triple IF staining of SOX17, TFAP2C, and the transcription factor PRDM1, all markers of early PGCs, identifies triple positive rPGCs in the dorsal mesentery as well as rPGCs closely associated with the gonadal ridge epithelium (Fig. 1d). Furthermore, all TFAP2C positive rPGCs in CS12 co-express the pluripotency transcription factor OCT4 (Fig. 1e). To determine whether rPGCs at CS12 correspond to early rPGCs (VASA negative) or late rPGCs (VASA positive) 6, we stained for VASA and the rPGC marker TFAP2C. Our data indicate that the majority of CS12 rPGCs were located in either the dorsal mesentery or the genital ridge epithelium, and the majority of TFAP2C positive rPGCs were also VASA positive (Supplementary Figure 1a, b). In addition, rare VASA negative TFAP2C positive cells were also identified (Fig. 1e, white arrows). As expected SOX2 was not expressed in any TFAP2C positive rPGCs (Fig. 1e).

Once we determined that the antibodies raised against TFAP2C, SOX17, PRDM1, OCT4, and VASA (Supplementary Table 1) could discriminate rPGCs in the embryo, we next examined the expression of these proteins with differentiation of riPSCs into rPGCLCs. The rPGCLC differentiation strategy involved a modification 5 of the two-step differentiation protocol first described by Sasaki and colleagues 5. The first step involves harvesting undifferentiated riPSCs cultured on mouse embryonic
fibroblasts (MEFs) as a single cell suspension followed by differentiating the riPSCs for 24h to create incipient mesoderm-like cells (iMeLCs) (Fig. 2, Supplementary Figure 2a). The second step involves differentiating the iMeLCs as three-dimensional aggregates in low adhesion 96-well plates (Fig. 2a). Rhesus PGCLCs are formed in step 2 within the aggregates in response to bone morphogenetic protein 4 (BMP4). In the first experiment, we used IF to test whether the transcription factors SOX17, PRDM1, or TFAP2C are expressed in undifferentiated riPSCs, or iMeLCs prior to aggregate differentiation. We also examined the expression of BRACHYURY (BRA) to confirm iMeLC induction, as well as pluripotency transcription factors OCT4 and SOX2. We found that both riPSCs and riMeLCs expressed OCT4 and TFAP2C, and in addition, riMeLCs also expressed BRA as expected2. SOX2 was uniformly expressed in riPSCs and heterogeneously in iMeLCs. Importantly, we found that riPSCs, and iMeLCs did not expressed SOX17 or PRDM1 (Supplementary Figure 2b, c). Therefore, we hypothesize these transcription factors could be used together to document the emergence of nascent rPGCLCs in the aggregate.

To identify rPGCLCs, aggregates were assessed at Day (D) 1 to D4 for SOX17 and PRDM1 expression. SOX17/PRDM1 double positive rPGCs. White arrowheads indicate TFAP2C positive VASA negative rPGCs. SOX2 (green) is not expressed in rPGCs. Merged images show DAPI nuclear stain (blue). Scale bars, 15 µm. n = 3 CS12 embryos, shown are images from Day (D) 28.
positive cells were identified starting on D1 through D4 with only a rare single positive cells (Supplementary Figure 3a). Previous studies analyzing cyno embryos at the time of lineage specification identified SOX17/PRDM1 double positive cells as marking both cynoPGCs as well as visceral endoderm (VE) cells. Therefore, to discriminate between these two possibilities in the aggregates, we also stained SOX17 together with TFAP2C, which is positive in cynoPGCs while being negative in VE. Double IF staining of aggregates revealed that subpopulations of SOX17 positive cells were also TFAP2C positive at D1, with more double positive cells emerging from D2 (Supplementary Figure 3b).

Based upon the SOX17/TFAP2C dual staining results, we examined rPGCLC formation by using the triple stain to identify cells co-expressing PRDM1/SOX17/TFAP2C in the aggregates. Using this approach, we discovered triple positive rPGCLCs as early as D1 of aggregate differentiation in two independent lines of riPSCs (riPSC89 and riPS90) (Fig. 2b and Supplementary Figure 3c). These triple positive rPGCLCs persisted through D4 and up to D8 (Fig. 2b and Supplementary Figure 3c). However, by D15 of aggregate differentiation, triple positive rPGCLCs were undetectable in either riPSC line (Supplementary Figure 3c–e).

Given that TFAP2C may be expressed earlier in the rhesus macaque relative to hPGCs, we quantified the total percent of TFAP2C positive cells from D1 to D4 of aggregate differentiation, and discovered that TFAP2C positive cells are a minor fraction of the total aggregate at each time point (Fig. 2c). Furthermore,
TFAP2C positive cells were mostly triple positive at all time points (TFAP2C together with SOX17 and PRDM1). However, TFAP2C/SOX17 double positive cells were clearly identified notably at D1–3 whereas TFAP2C/PRDM1 double positive cells were never identified (Fig. 2d). Taken together, our results indicate that triple positive PRDM1/SOX17/TFAP2C rPGCLCs are induced at D1 of aggregate differentiation, and that the potential rPGCLC precursors are either SOX17/TFAP2C or SOX17/PRDM1 double positive but not TFAP2C/PRDM1.

**Rhesus macaque PGCLCs arrest before epigenetic reprogramming.** To identify the transcriptional identity of rPGCLCs, we isolated rPGCLCs using fluorescence activated cell sorting (FACS) at D1, 2, 4, and 8 of aggregate differentiation using antibodies that recognize EPCAM and ITGA6 (Fig. 3a). EPCAM and ITGA6 were chosen because they were previously used to recognize hPGCLCs in aggregate differentiation. The frequency of EPCAM/ITGA6 double positive cells was highest in D2 aggregates, while at D4, this population stabilized at around 9.5%. On D8, the percent of EPCAM/ITGA6 ultimately constitutes only a small fraction of all cells in an aggregate (Fig. 3a). Using real-time polymerase chain reaction (PCR) to detect rPGC markers NANS3, TFAP2C, PRDM1, and cKIT, we show that the rPGCLC genes are all up-regulated in the EPCAM/ITGA6 double positive population compared to undifferentiated r iPSCs and rMeLCs (Fig. 3b).

To determine whether rPGCLCs are in the early or late stages of rPGC development, we examined VASA expression using IF at D4 and D8 and discovered that all TFAP2C positive rPGCLCs in aggregates are VASA negative (Fig. 3c). In contrast, late rPGCs at D4 and D8 and discovered that all TFAP2C positive rPGCLCs in aggregates were all positive for 5mC (Fig. 3d). This finding that xenotransplantation supported rPGCLC differentiation, sections of testicles were processed to paraffin and sections stained for VASA together with the spermatogonial stem cell markers MAGEA4 and ENO2. Using these markers, we show that almost all Nhp positive rPGCLCs in the recipient mouse testicular tubules differentiate into VASA positive germ-line cells confirming the results of the whole mount. Furthermore, these Nhp-positive germ line cells were exclusively located on the basement membrane, typical of spermatogonia. A fraction (28%) of NHP positive rPGCLCs also expressed the spermatogonial protein MAGEA4 (Fig. 4b). In contrast, ENO2 was not expressed in any Nhp-positive cells (n=15) in the mouse testis (Fig. 4b).

**Xeno- and homologous transplantation promote differentiation.** Previously, it was reported that CS23 rPGCs xeno-transplanted into the seminiferous tubules of the busulfan-treated mouse testicles colonize the seminiferous tubule basement membrane giving rise to colony-like chains of cells. To determine whether rPGCLCs persist in this assay and form colonies, we first created a green fluorescent protein (GFP) expressing subline of r iPSCs using a lentivirus (referred to as r iPSC89UbC-GFP). Prior to transplantation, karyotype analysis was performed to confirm a normal 42XX male karyotype. The r iPSC89UbC-GFP line was differentiated through the two-step protocol ending at D8 of aggregate differentiation. The D8 aggregates were shipped overnight on cold packs to MD Anderson Cancer Center in Houston where they were dissociated and xenotransplanted into the seminiferous tubules of adult immune-deficient nude mice that had been irradiated to ablate endogenous spermatogenesis. The recipient mouse testes were seeded with between 5.0 × 10^5 and 1.3 × 10^5, unsorted aggregate cells per transplant (n=6 testicles). As a control, we also transplanted 1.9 × 10^5–2.7 × 10^5 undifferentiated r iPSCs per transplant (n=6 testicles). Eight weeks after transplant the mice were euthanized, and the testicles were analyzed.

At the time of dissection, 6/6 recipient testicles receiving a single cell suspension of aggregate cells were GFP positive. In contrast, only 4/6 recipient testicles that received undifferentiated r iPSCs were GFP positive (Supplementary Figure 4a). This result is consistent with previous reports that single cell suspensions of undifferentiated primate pluripotent stem cells exhibit poor survival, which our data would suggest extends to the seminiferous tubule environment. Whole mount IF of the GFP positive testicles using the non-human primate (Nhp), rhesus macaque testis cell antibodies followed by AlexaFluor488 secondary antibodies demonstrated that the Nhp/GFP positive cells in the recipient testicles that received aggregates were VASA positive indicative of rPGCLC differentiation. In contrast, the recipient testicles that received undifferentiated r iPSCs were VASA negative (Fig. 4a). Histological analysis of the GFP positive recipient testicles xenotransplanted with donor aggregates sometimes revealed cysts (4/6) (Supplementary Figure 4b–e, asterisk) but, did not form teratomas. In contrast, testicles transplanted with r iPSCs yielded teratomas in 2 out 4 of cases.

To confirm that xenotransplantation supported rPGCLC differentiation, sections of testicles were processed to paraffin and sections stained for VASA together with the spermatogonial stem cell markers MAGEA4 and ENO2. Using these markers, we show that almost all Nhp positive rPGCLCs in the recipient mouse testicular tubules differentiate into VASA positive germ-line cells confirming the results of the whole mount. Furthermore, these Nhp-positive germ line cells were exclusively located on the basement membrane, typical of spermatogonia. A fraction (28%) of NHP positive rPGCLCs also expressed the spermatogonial protein MAGEA4 (Fig. 4b). In contrast, ENO2 was not expressed in any Nhp-positive cells (n=15) in the mouse testis (Fig. 4b).

Due to the finding that only one of the spermatogonial stem cell markers (MAGEA4) was expressed in xenotransplanted donor rPGCLCs, we hypothesized that the rPGCLCs may not correspond to adult-stage spermatogonia and instead the MAGEA4 single positive germ cells may represent an earlier stage of primordial germ cell development. To address this, we stained human fetal testis at 17 weeks post-fertilization and found ENO2/MAGEA4 double positive hPGCs (hPGCs were identified with VASA) (Supplementary Figure 4f). In contrast, at earlier embryonic points (D28 and D50), MAGEA4 was expressed in VASA positive rPGCs whereas ENO2 was not (Supplementary Figure 4g). Given that MAGEA4 and ENO2 were not expressed in rPGCLCs before xenotransplantation (Supplementary Figure 4g), our results support the hypothesis that xenotransplantation results in rPGCLC differentiation into VASA/MAGEA4 positive rPGCLCs corresponding to a stage in embryo development prior to ENO2 positive spermatogonial differentiation.

To confirm that the rPGCLCs originated from the specified rPGCLCs in the aggregates and not de novo specification of rPGCLCs donor aggregate cells while in the seminiferous tubules, we used FACS to isolate EPCAM/ITGA6 positive (r PGCLCs) and negative (somatic) cells, and xenotransplanted the two populations separately into the testicles of irradiated SCID mice (Supplementary Figure 4h). Using confocal microscopy of intact testicles, we identified GFP-positive signal only in the testicles xenotransplanted...
with FACS isolated rPGCLCs but not in testicles xenotransplanted with somatic cells or sham control (no donor cells) (Supplementary Figure 4h). This result suggests that the rPGCLCs in the xenotransplant differentiated from the rPGCLCs originally specified in the aggregates, which subsequently underwent further differentiation following xenotransplantation.

Based on the induction of MAGEA4 and VASA in the xenotransplanted donor cells, the expectation is that the colonized rPGCLCs should also have initiated epigenetic reprogramming. To address this, we transplanted a single cell suspension of D4 aggregates from the riPSC90 line into busulfan-treated nude mice and examined enrichment of 5mC and 5hmC using IF. Given that...
riPSC90 was not labeled with GFP, the donor rPGCLCs were identified using an antibody that detects the primate-specific, nuclear mitotic apparatus protein (NuMA). This result shows that the NuMA positive rPGCLCs undergo global loss of 5mC (Fig. 5a). To detect 5hmC in these xenotransplants, we performed sequential IF staining of adjacent sections, and show that NuMA positive rPGCLCs are also 5hmC positive (Supplementary Figure 5a–d). Therefore, this result supports the hypothesis that xenotransplantation induces rPGCLC differentiation to the equivalent of late-stage rPGCs in the middle of epigenetic reprogramming similar to those identified at CS12, but younger than those at CS23 where 5hmC and 5mC are no longer detectable by IF.

Given that the loss of 5hmC in mPGCs occurs largely through replication-coupled demethylation, we were also interested in determining whether rPGCs are in cycle during this developmental window. To achieve this we stained for the nuclear antigen Ki67 in rhesus embryos at CS12 (D28) and CS23 (D50) as well as aggregates at D4 and D8 (Fig. 5b). In all cases, the rPGCLCs/rPGCs were identified using IF for TFAP2C (red) with Ki67 labeled in green. The results are quantified as the percentage of Ki67 positive cells in the TFAP2C positive or negative population (Fig. 5c). During aggregate differentiation, the TFAP2C positive rPGCLCs at D4 are mostly Ki67 negative (not in cycle), whereas the somatic cells are Ki67 positive. In contrast

---

**Fig. 4** Xenotransplantation of rPGCLCs leads to expression of late stage markers. **a** Whole mount immunofluorescence (IF) staining of mouse testicle seminiferous tubules xenotransplanted with a single cell suspension of Day (D) D8 riPSC90UbC::GFP aggregates, reveals the presence of Nhp/GFP (green) cells, near the basement membrane of the tubules (white dotted outline) that co-express the late-stage rPGC marker VASA (red). Recipient testicles xenotransplanted with riPSCs contained Nhp/GFP positive cells but did not express VASA. Scale bars, 40 µm. **b** IF staining on paraffin-embedded, testicle sections from recipients that received a single cell suspension of aggregate cells. Nhp-positive cells (green) are found inside the seminiferous tubules (basement membrane, white dotted outline) and express VASA (white arrows). A subset of Nhp-positive cells expressed MAGEA4 (red, white arrows), while ENO2 was absent. Scale bars, 40 µm. **c** Positive control staining for VASA (red), MAGEA4 (red), and ENO2 (red) in adult rhesus macaque testis. Scale bars, 40 µm.
at D8, the majority of somatic cells are Ki67 negative, whereas 50% of the rPGCLCs are now in cycle and are Ki67 positive. In the embryo, rPGCs at D28 and D50 are almost Ki67 positive and therefore in the cell cycle during 5mC/5hmC epigenetic reprogramming. Analysis of the xenotransplants shows that the majority of NuMA positive donor cells (red) are Ki67 negative and therefore the rPGCLCs are not in cycle following xenotransplantation (Fig. 5d). Taken together, our data indicate that xenotransplantation induces rPGCLC differentiation and epigenetic remodeling, however the adult testicular niche causes the rPGCLCs to exit the cell cycle before completing epigenetic reprogramming and global removal of 5hmC. Due to the inability of primate cells to fully reprogram and differentiate to ENO2 positive spermatogonia in the mouse testicular niche, we next tested the hypothesis that homologous transplantation of aggregates into irradiated rhesus macaque male testicles could remove the barrier to rPGCLC differentiation. Therefore, we transplanted dissociated aggregates containing rPGCLCs or undifferentiated riPSCs into the left and right testis, respectively, of rhesus macaque males depleted of germ cells by irradiation (n = 2). The monkeys were also administered GnRH-antagonist treatment starting immediately after irradiation for 2 months until the time of transplantation. GnRH-antagonist treatment is expected to facilitate survival and colonization of transplanted cells16. Serum testosterone levels were monitored to confirm transient suppression of testosterone during GnRH antagonist treatment and maintenance of Leydig cell function after irradiation (Fig. 6a). Ultrasound-guided rete testis injection was performed under general anesthesia to transplant 4.2 million single aggregate cells, and 9.8 million undifferentiated riPSCs into the rete testis. Following transplantation, testicular volumes and serum testosterone were measured monthly in each animal.

Fig. 5 Xenotransplantation is associated with rPGCLCs epigenetic reprogramming and precocious exit from the cell cycle. a Donor cells are detected through their expression of the nuclear mitotic apparatus protein (NuMA, red) in paraffin sections of recipient testicles (n = 13) xenotransplanted with a single cell suspension of Day (D) D4 aggregate cells derived from riPSC90. 5mC (green) was detected in recipient somatic cells. Scale bar, 50 µm. b IF on paraffin sections of aggregates at D4 and D8, and rhesus embryos at CS12 (D28) and CS23 (D50) for the rPGCLC/rPGC marker TFAP2C (red) and the proliferation marker, Ki67(green). Nuclei were detected using DAPI (blue). Scale bars, 50 µm. c Quantification of Ki67 positive (blue bar graphs) and Ki67 negative (gray bar graphs) rPGCLCs/rPGCs and somatic cells in aggregates at D4 and D8, and rhesus embryos at CS12 (D28) and CS23 (D50). All data points used to generate bar graphs (black circles) were overlaid to the corresponding bar graphs, error bars represent the Standard Error of the Mean (S.E. M.), N = 3 technical replicates. d Recipient mouse testicles (n = 13) xenotransplanted with riPSC90 D4 aggregate cells co-stained for NuMA (red) and Ki67 (green). Nuclei were detected using DAPI (blue). Scale bar, 50 µm.
Fig. 6 Homologous transplant of rPGLCs results in differentiation to late rPGC stage. a The volume of recipient testicles that received donor Day (D) D8 riPSC89UbiC:GFP aggregate cells returned to baseline levels, comparable to the volume prior to irradiation and subsequent GnRH antagonist treatment. In contrast one recipient male who received riPSC89UbiC:GFP cells exhibited a volume increased. b H&E staining of the abnormally large recipient testicle transplanted with riPSC89UbiC:GFP revealed teratomas. Scale bars, 400 µm. c H&E staining of aggregate recipient testicles reveals large cysts (asterisks) near normal tubules. Scale bars, 400 µm. d Expression of GFP in recipient testicles transplanted with D8 riPSC89UbiC:GFP aggregate cells within the seminiferous tubules (white dotted lines). Scale bars, 40 µm. e A graph of the number of transplants that had GFP positive donor cells. f Graph of the number of GFP positive transplants with colonies. g–j IF on sections of rhesus recipient testicles that received a single cell suspension of aggregate donor cells (GFP positive, green). The surviving donor cells co-expressed the late stage rPGC markers vASA (red) and MAGEA4 (red), but not ENO2. Nuclear staining using DAPI (blue) is shown in merged panels for g–i. Scale bars, 40 µm.
Here, we establish the rhesus macaque as another non-human primate species amendable to rPGCLC differentiation in vitro. Consistent with previous studies in the human and in the cynomolgus macaque, VASA RNA and protein are not detected in rPGCLCs during aggregate formation. VASA is an evolutionary conserved gene that is expressed in germline cells of all metazoan, and in mammals is expressed as PGCs approach the developing genital ridge. In the current study, we show that a genital ridge is not necessary to induce VASA protein expression in rPGCLCs, and instead VASA protein can be induced following xenogeneic transplantation into an adult testicular niche.

The mechanism responsible for inducing VASA protein in mammalian PGCs is unclear. Using transgenic mouse technology, ectopic expression of mouse vasa homologue (mvh) in somatic cells can be achieved by the global loss of DNA methylation. In the current study, we show that VASA-positive rPGCLCs is accompanied by a global loss of 5mC. Therefore, we hypothesize that the expression of VASA in rPGCLCs is associated with DNA methylation reprogramming following xenotransplantation. Given that rPGCLCs are lost between D8 and D15 of aggregate differentiation in vitro, we propose that it is the transfer of rPGCLCs to a new environment supportive of rPGCLC survival that is primarily responsible for triggering global DNA demethylation and advancing differentiation into VASA-positive cells. Consistent with this, the culture of mPGCs/mPGCLCs in growth factors and chemicals that maintain mPGC survival also enable the transition to MVH-positive state without a requirement for testicular cells. Therefore, it is likely that adult testicular cells per se are not responsible for inducing VASA expression. Given this result, we anticipate that chemically defined conditions could be identified in the future to support the differentiation of rPGCLCs without the requirement for transplantation.

The markers MAGEA4 and ENO2 were originally chosen as an approach to examine rPGCLC differentiation to adult spermatogonial stem cells following xenotransplantation. Given the lack of ENO2 expression in MAGEA4 positive transplant-derived rPGCLCs, we hypothesized that MAGEA4 may be marking an earlier stage of germ cell differentiation when ENO2 is negative. Here, we show that MAGEA4 is expressed much earlier in rhesus germ cell development, being present in VASA positive rPGCs at CS12 as they approach the genital ridge epithelium, as well as in VASA positive rPGCs at CS23. In contrast ENO2 is not expressed in rPGCs at these stages, instead being expressed together with MAGEA4 and VASA in putative pre-spermatogonia. It is difficult to predict whether the lack of ENO2 in the transplant-derived rPGCLCs is due to a block in developmental timing, or whether ENO2 is abnormally silenced. However, failure to also erase 5hmC in the xenotransplant-derived rPGCLCs, together with exit from the cell cycle is consistent with a block in developmental timing such that the transplanted rPGCLCs are arrested in development and exiting the cell cycle at a stage equivalent to rPGCs younger than CS23.

Previous studies showed that testicular xenotransplantation of undifferentiated hESCs and hiPSCs into the testicles of busulfan-treated nude mice causes differentiation into VASA positive rPGCLCs accompanied by teratoma, embryonal carcinoma, and yolk sac tumor formation. The tumors in previously published studies most likely originate from hiPSCs/hESCs because testicular xenotransplantation of embryonic testicles containing bona fide hPGCs results in colony formation without tumorigenesis. Unlike studies using human pluripotent stem cells, we did not identify any evidence of VASA positive putative rPGCLCs in the testicles transplanted with undifferentiated rPSCs. The underlying reason for this discrepancy is unclear and may be due to species-specific differences in the donor cells (human versus rhesus).

Successful resumption of spermatogenesis from donor cells in homologous monkey-to-monkey transplants involves the transfer of tissues containing adult spermatagonia. In the current study, the lack of spermatogenesis in homologous transplants could be due to either an incompetent rPGCLC or alternatively the possibility that rPGCLC/rPGCs are incapable of differentiating in an adult niche. In support of the young versus old niche hypothesis, mouse transplantation studies using neonatal infertile recipients can support spermatogenesis from mPGCs, whereas adult infertile recipients can support spermatogenesis from pre-spermatogonia or spermatogonia. Given this, the only way to experimentally test whether rPGCs have the competency to fully differentiate into haploid sperm in an adult niche would be to perform homologous transplants of bona fide rPGCs into infertile rhesus recipients at different ages (from neonate to adult). The feasibility of this theoretical experiment is challenged by the very low numbers of rPGCs found during the embryonic stages of interest.

In summary, we show that rPGCLCs differentiated in aggregates are competent to differentiate into VASA/MAGEA4...
Methods

Time-mated breeding of rhesus macaques. Time-mated breeding of rhesus macaques and females was performed using estradiol大胆的 from the female starting from D5 to D8 after menses began27. Pregnancy was confirmed by measuring progesterone as well as by ultrasound. Embryos at CS21 were selected by C-section (n = 3 at D28, n = 1 at D26, n = 1 at D30). All rhesus macaque time-mated breeding experiments were conducted following the approval of the Oregon National Primate Research Center (ONPRC) Institutional Animal Care and Use Committee (IACUC).

Maintenance and directed differentiation of riPSC lines. Undifferentiated riPSC8927 and riPSC9029 cells were cultured on a feeder layer of mitomycin C-treated MEFs in hESC media (DMEM/F-12) (Life Technologies), 20% KSR (Life Technologies), 10 ng/mL bFGF (R&D Systems), 1% nonessential amino acids (Life Technologies), 2 mM L-glutamine (Life Technologies), Primocin® (InvivoGen), and 0.1 mM β-mercaptoethanol (Sigma). Media was changed daily and colonies were passaged manually every 5 days. For iMELC inductions we followed previously published protocols with minor modifications23,24. Briefly, D5 riPSCs colonies were trypsinized (0.05% trypsin) (Life Technologies), resuspended in MEF media. The MEFs were depleted by plating the cell suspension in tissue culture dishes, two times, for 5 min each. The resulting cell suspension was pelleted and resuspended in iMELC media (GMEM) (Life Technologies), 15% KSR (Life Technologies), 0.1 mM nonessential amino acids (Life Technologies), penicillin/streptomycin/L-glutamine (Life Technologies), Primocin® (InvivoGen), 0.1 mM β-mercaptoethanol (Sigma), sodium pyruvate (Life Technologies), activin A (PeproTech), CHIR99021 (Stemgent), Y-27632 (Stemgent), 10 mM insulin, 10 mM follicle stimulating hormone (FSH), 10 mM human chorionic gonadotropin, and plated at a density of 1.0 × 10^6 cells per well of a human plasma fibronectin (Invitrogen)-coated 12-well plate. After 24 h of incubation at 37 °C with 5.0% CO2, iMECs were trypsinized (0.05% trypsin, Life Technologies) and resuspended in PGCCLC media (GMEM) (Life Technologies), 15% KSR (Life Technologies), 0.1 mM nonessential amino acids (Life Technologies), penicillin/streptomycin/L-glutamine (Life Technologies), Primocin® (InvivoGen), and 0.1 mM β-mercaptoethanol (Sigma), sodium pyruvate (Life Technologies), 10 ng/mL human LIF (EMD Milipore), 200 ng/mL BMPA (R&D Systems), 50 ng/mL EGF (Fisher Scientific), 10 µM Y-27632 (Stemgent), and plated at a density of 1.0 × 10^5 cells per well of a low adherence spheroid forming 96-well plate (Corning). Aggregates were collected for analysis on Days 1, 2, 3, 4, 8, and 15 of directed PGCCLC differentiation.

Human fetal samples. Human prenatal tests were acquired following elected termination and pathological evaluation after UCLA-IRB review which deemed the project exempt under 45 CRF 46.102(f). All prenatal tests were obtained from the University of Washington Birth Defects Research Laboratory (BDRL), under the regular oversight of the University of Washington IRB approved Human Subjects protocol combined with a Certificate of Confidentiality from the Federal Government. BDRL collected the fetal testes and shipped them overnight in HBSS with an ice pack for immediate processing in Los Angeles. All consented material was donated anonymously and carried no personal identifiers. Developmental age was documented by BDRL as days post-fertilization using a combination of pre-natal intakes, foot length, Streater’s Stages, and crown-rump length. All prenatal gonads documented with a birth defect or chromosomal abnormality were excluded from this study.

Immunofluorescence (IF) staining. Aggregates containing rPGCLCs were collected and then fixed in 4% PFA and embedded in histogel (Thermo Scientific) to facilitate subsequent embedding into paraffin blocks. Sections of aggregates, rhesus embryo, or human fetal tissues (5 µm) placed onto microscope slides were then de-paraffinized and rehydrated through a xylene, ethanol series. For antigen retrieval, slides were heated to 95 °C in Tris-EDTA solution (10 mM Tris Base, 1 mM EDTA solution, 0.05% Tween-20, pH 9.0). Sections were permeabilized (PBS, 0.05% Triton-100) and then blocked in PBS containing 10% normal donkey serum and 0.3% Triton-X (AM900210; Life Technologies). Anti-5hmC (91185S; 1:100), anti-Oct4 (sc262; 1:100), anti-OCT3/4 (sc5279; 1:100), anti-TFAP2C (sc21762; 1:200), anti-TFAP2C (sc8977; 1:200), anti-PRDM1 (9115S; 1:100), anti-SOX17 (GT15094; 1:100), anti-SOX2 (MAB2018; 1:100), anti-VASA (AF2030; 1:100), anti-Brachyury (Chrom 57R; 1:30), anti-Brachyury (Chrom 57R; 1:30), anti-EN02 (MM5-518P; 1:100), and anti-SOX3 (556003; 1:200), anti-NuMa (AB48680; 1:200) (Supplementary Table 1). Monoclonal antibodies (Prolong-gold anti-fade w/DAPI, Invitrogen) were added and slides were sealed. Slides were allowed to cure for at least 24 h at 4 °C prior to imaging. riPSCs and rMeLCs were plated onto culture slides (Corning) and grown overnight their respective media at 37 °C with 5% CO2. Cells were washed with PBS then fixed in 4% PFA for 10 min. Cells were rinsed with PBST and then permeabilized (PBS, 0.05% Triton-100) for 10 min. Nonspecific binding was blocked (PBS, 10% normal donkey serum for 30 min) and then incubated overnight at 4 °C with primary antibodies (Supplementary Table 1). Species-specific secondary antibodies were added to slides and incubated at room temperature for 1 h. Mounting media was added, slides were sealed, and then cured for 24 h.

Mouse testes tissues were fixed overnight in 4% paraformaldehyde, 5 µm thick paraffin sections were used for the immunofluorescence study. The monkey testes were fixed in cold 4% paraformaldehyde overnight, processed through a sucrose gradient, and then embedded in low-iodine compound: 80% sucrose, 20% paraformaldehyde for IF staining. For testicular histology of both mouse and monkey, Bouin’s or 4% paraformaldehyde-fixed tissues were embedded in paraffin and sections were stained in periodic acid-Shiff’s-hematoxylin.

For fluorescent immunostaining, either the deparaffinized or frozen sections were subjected to antigen retrieval by initially heating in boiling citrate buffer (BioGenex) for 2 min followed by cooling for 30–60 min. The slides were rinsed and blocked in antibody diluent. Subsequently, sections were stained with the following primary antibodies in antibody diluent: goat anti-VASA (AF2030; 1:100), mouse anti-EN02 (MM5-518P; 1:500), mouse anti-MAGE-A4 (Clone 57R, kindly provided by Dr. Giulio Spagnoli, University Hospital Basel, Switzerland; 1:30), rabbit anti-GFP (29565; 1:100), rabbit anti-rhesus testis cell (NHP) (provided by Kyle Orwig; 1:200). For double immunofluorescent staining, the two primary antibodies incubated with the tissue sections were detected with species-specific secondary antibodies (Supplementary Table 1). Stained sections were mounted with VectaShield mounting medium containing DAPI (Vector Laboratories) and imaged. Positive immunoreactivity was validated by the omission of primary antibody and testing the antibody in a tissue in which it is known to be positive.

Microscopy. Confocal images of the riPSC89 line, riPSC900 line, iMeLCs, and sectioned embryos were examined on an LSM 880 (Carl Zeiss) with a Plan-Apochromat 20×/0.8 NA and a Plan-Apochromat 63×/1.4 NA M27 oil immersion objective at room temperature. Acquired images were processed using IMARIS 8.1 (Bitplane). Hi&El slides were examined on an Olympus BX-61 light microscope. Images were processed with Image J version 1.51d (NIH). D28 embryos, teratomas, and cysts images were stitched together using the grid-stitching plugin in Image J version 1.51d (NIH).

For immunofluorescence imaging of tissue sections or whole mounts of seminiferous tubules, a Leica DM 4000B (Leica Microsystems) microscope was used.

Image analysis. Quantification of confocal images was performed using IMARIS 8.1 (Bitplane) microscopy image analysis software. To quantify the total percent of cells that were positive for TFAP2C in D1, D2, D3, and D4 aggregates, we first quantified the total number of nuclei using the spot detection function to detect DAPI. Only DAPI positive nuclei that overlapped with TFAP2C signals were counted, using the IMARIS spot function. Using the TFAP2C channel and the colocalization function in IMARIS we built co-localization channels for SOX17 and PRDM1 channels. Through this strategy, we were able to quantify single, double, and triple positive cells found in aggregates using the IMARIS spot function. Three random aggregates were counted for each experiment, and for Days 1–4 of the co-localization experiments were repeated on two separate cell lines (riPSC89 and riPSC90) with at least 2 technical replicates each. Graphs were made using Prism 7 (GraphPad) data analysis software.

To quantify the total percent of cells that were positive (+) for Ki67, first we counted the total number of nuclei using the spot detection function to detect DAPI. Next, utilizing TFAP2C to demarc germ cells we built a co-localization channel for Ki67. The spot function was used on the newly built co-localization
channel to quantify the number of cells that were both positive for TFAF2C and Ki67. To quantify the number of somatic cells that were Ki67+ we counted the total number of cells that were positive for Ki67 and subtracted from the total number of Ki67+ germ cells (TFAF2C+). Three random fields were counted for each sample with 3 replicates each for Days 1–4 of differentiation. Graphs were made on Prism 7 (Graph Pad) and error bars represent the standard error of the mean.

Fluorescence activated cell sorting analysis of rPGCLCs. The rPGCLCs at D1, D2, and D4 were dissociated using 0.05% trypsin, while D8 PGCLCs were dissociated using 0.05% trypsin and Collagenase IV. Cells were washed with MEF media and then re-suspended in FACS buffer. Dissociated cells were incubated with anti-ITGA6-BV421 (BioLegend 136234; 1:50) and EPACM-PE (Life Technologies A15782; 1:50) antibodies. Double positive cells were collected using an ARIA-H Fluorescence Activated Cell Sorter. Cells were sorted into RLT buffer and stored at −20 °C until ready to isolate RNA. Sorts of rPGCLCs were performed on at least 2 replicates for each group. Cytometry analysis was performed using FlowJo version 10.

Establishment of the riPSC89UbiC:GFP reporter line. The riPSC89 GFP reporter line was established through episomal delivery of the lentivirus GFP-IRES_PUR-O cassette, a kind gift from Dr. Zoran, using lipofectamine 3000 (Invitrogen). Karyotypes performed by Cell Line Genetics (Madison, WI) indicated that the GFP-expressing riPSC89 line had a normal karyotype.

Whole mount immunofluorescence of seminiferous tubules. For quantitative analysis of donor rhesus testis cell colonization, intact seminiferous tubules were prepared from nude mouse recipient testes, collected 8 weeks after transplantation16. Donor-derived colonies of spermatogonia were detected in intact seminiferous tubules by whole mount immunofluorescent staining with the rhesus testis cell antiserum18. Samples were dehydrated stepwise in methanol and then incubated in MeOH:DMSO:H2O2 (4:1:1) for 2 h. The rhesus testis-cell antibody was used and detected with rabbit secondary antibodies (Supplementary Table 1). Samples were mounted with Vectashield medium containing DAPI (Vector Laboratories). All relevant data are included in the paper and/or its Supplementary Information. Data availability

Received: 7 November 2017 Accepted: 16 November 2018

Published online: 17 December 2018

References

1. Irie, N. et al. SOX17 is a critical specifier of human primordial germ cell fate. Cell 160, 253–268 (2015).
2. Sasaki, K. et al. Robust in vitro induction of human germ cell fate from pluripotent stem cells. Cell Stem Cell 17, 178–194 (2015).
3. Kosjima, Y. et al. Evolutionarily distinctive transcriptional and signaling programs drive human germ cell lineage specification from pluripotent stem cells. Cell Stem Cell 21, 517–532 (2017).
4. Kobayashi, T. et al. Principles of early human development and germ cell program from conserved model systems. Nature 546, 416–420 (2017).
5. Chen, D. et al. Germline competency of human embryonic stem cells depends on comodoserin. Biol. Reprod. 9, 850–861 (2017).
6. Sasaki, K. et al. The germ cell fate of cynomolgus monkeys is specified in the nascent amnion. Dev. Cell 39, 169–185 (2016).
7. Jung, D. et al. In vitro differentiation of human embryonic stem cells into ovarian follicle-like cells. Nat. Commun. 8, 15680 (2017).
8. Nakaki, F. et al. Induction of mouse germ-cell fate by transcription factors in vitro. *Nature* **501**, 222–226 (2013).
9. Hayashi, K., Ohba, H., Kurimoto, K., Arakami, S. & Saitou, M. Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* **146**, 519–532 (2011).
10. Hirota, T. et al. Fertile offspring from sterile sex chromosome trisomic mice. *Science* **357**, 932–935 (2017).
11. Zhou, Q. et al. Complete reprogramming from embryonic stem cell-derived germ cells in vitro. *Cell Stem Cell* **18**, 330–340 (2016).
12. Nagan, M., Patrizio, P. & Brinster, R. Long-term survival of human spermatogonial stem cells in mouse testes. *Fertil. Steril.* **78**, 1225–1233 (2002).
13. Valli, H. et al. Fluorescence- and magnetic-activated cell sorting strategies to isolate and enrich human spermatogonial stem cells. *Fertil. Steril.* **102**, 566–580 (2019).
14. Izadyar, F. et al. Identification and characterization of repopulating spermatogonial stem cells from the adult human testis. *Hum. Reprod.* **26**, 1296–1306 (2011).
15. Maki, C. et al. Phenotypic and molecular characterization of spermatogonial stem cells in adult primate testes. *Hum. Reprod.* **6**, 1480–1491 (2009).
16. Shetty, G. et al. Hormone suppression with GnRH antagonist promotes spermatogenic recovery from transplanted spermatogonial stem cells in irradiated cynomolgus monkeys. *Andrology* **1**, 886–898 (2013).
17. Clark, A. T. et al. Primate primordial germ cells acquire transplantation potential by Carnegie stage 23. *Stem Cell Rep.* **9**, 329–341 (2017).
18. Hermann, B. P. et al. Characterization, cryopreservation, and ablution of spermatogonial stem cells in adult rhesus macaques. *Stem Cells* **25**, 2330–2338 (2007).
19. Hermann, B. P. et al. Molecular dissection of the male germ cell lineage identifies putative spermatogonial stem cells in rhesus macaques. *Hum. Reprod.* **24**, 1704–1716 (2009).
20. Hermann, B. P. et al. Spermatogonial stem cell transplantation into rhesus testes regenerates spermatogenesis producing functional sperm. *Cell Stem Cell* **11**, 715–726 (2012).
21. Clarke, R. L. et al. The expression of Sox17 identifies and regulates haemogenic endothelium. *Nat. Cell Biol.* **15**, 502–510 (2013).
22. Seisenberger, S. et al. The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. *Mol. Cell* **48**, 849–862 (2012).
23. Hackett, J. A. et al. Germine DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. *Science* **339**, 448–452 (2013).
24. Yamaguchi, S. et al. Dynamics of 5-methylcytosine and 5-hydroxymethylcytosine during germ cell reprogramming. *Cell Res.* **23**, 329–339 (2013).
25. Calvopina, J. H. et al. The aorta-gonad-mesonephros organ culture recapitulates ShmC reorganization and replication-dependent and independent loss of DNA methylation in the germline. *Stem Cells Dev.* **24**, 1536–1545 (2015).
26. Sosa, E. et al. An integration-free, virus-free rhesus macaque induced pluripotent stem cell line (rIPSC9) from embryonic fibroblasts. *Stem Cell Res.* **17**, 444–447 (2016).
27. Okumura, N. et al. Enhancement on primate corneal endothelial cell survival in vitro by a ROCK inhibitor. *Invest. Ophthalmol. Vis. Sci.* **50**, 3680–3687 (2009).
28. Mochizuki, K., Nishimia-Fujisawa, C. & Fujisawa, T. Universal occurrence of the vasa-related genes among metazoans and their germline expression in *Hydra*. *Dev. Genes Evol.* **211**, 299–301 (2001).
29. Castrillon, D. H., Quade, B. J., Wang, T. Y., Quigley, C. & Crum, C. P. The human vasa gene is specifically expressed in the germ cell lineage. *Proc. Natl. Acad. Sci. USA* **97**, 9585–9590 (2000).
30. Maoutok, D. M. et al. DNA methylation is a primary mechanism for silencing postmitochondrial primordial germ cells in both germ cell and somatic cell lineages. *Development* **133**, 3411–3418 (2006).
31. Oliveros-Etter, M. et al. PGd reversion to pluripotency involves erosion of DNA methylation from imprinting control centers followed by locus-specific re-methylation. *Stem Cell Rep.* **5**, 337–349 (2015).
32. Obia, H. et al. In vitro expansion of mouse primordial germ cell-like cells recapitulates an epigenetic blank slate. *EMBO J.* **36**, 1888–1907 (2017).
33. Durruthy, D. J. et al. Fate of induced pluripotent stem cell following transplantation to murine seminiferous tubules. *Hum. Mol. Genet.* **23**, 3071–3084 (2014).
34. Sosa, E. et al. An integration-free, virus-free rhesus macaque induced pluripotent stem cell line (rIPSC9) from embryonic fibroblasts. *Stem Cell Res.* **21**, 5–8 (2017).
35. Meistrich, M. L., Hunter, N. R., Suzuki, N., Tростle, P. K. & Withers, H. R. Gradual regeneration of mouse testicular stem cells after exposure to ionizing radiation. *Radiat. Res.* **74**, 349–362 (1978).
36. Zhang, Z., Shao, S. & Meistrich, M. L. The radiation-induced block in spermatogonial differentiation is due to damage to the somatic environment, not the germ cells. *J. Cell Physiol.* **211**, 149–158 (2007).

**Acknowledgements**

This project was funded by support from PO1HD075795 (K.E.O., A.T.C., G.S., M.L.M.). Rhesus CS12 embryo collections were supported by the grant P51 OD011092 (J.D.H. and A.T.C.). E.S. acknowledges the support of the Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research at UCLA Training Program for Post-doctoral Fellows. We would also like to acknowledge the BSCRC FACS Core and the BSCRC Imaging Core. We would also like to acknowledge Them T. Phan for help with the xenotransplantation experiments. E.J.R. was supported by the National Institute of General Medical Sciences of the National Institutes of Health under award number R25GM055652 awarded to T. Hasson. Human fetal tissue research is supported by a grant to Ian Glass at the University of Washington Birth Defects laboratory 5R24HD008363–53.

**Author contributions**

E.S. conceived and performed experiments and wrote the manuscript. D.C. performed FACS on rPGCs and rPGCCls and prepared cells for transplant. E.J.R. performed immunofluorescence experiments on rPGCs and rPGCCls. M.S., Z.W., and T.N.L. performed xenotransplantation experiments. K.A.P. and J.M.M. helped with homologous transplantation in rhesus macaques. R.C.T. performed dose and radiation field planning on each individual monkey. J.D.H. conceived experiments, maintained rhesus macaque IACUC approval for time-mated breedings, and oversight of rhesus macaque work in Oregon. M.L.M. performed experiments, maintained IACUC approval, and oversight of mouse xenotransplantation and rhesus macaque transplantation in Houston. K.E.O. conceived the experiments, performed homologous rhesus macaque transplants. G.S. conceived and performed experiments, maintained IACUC approval, and oversight of mouse xenotransplantation and rhesus macaque transplantation in Houston. K.E.O. conceived the experiments, performed experiments, maintained Institutional Biosafety Approval for rhesus macaque rIPSCs/rPGCCls work and wrote the manuscript.

**Additional information**

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-07740-7.

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

**Reprints and permission** information is available online at http://npg.nature.com/reprintsandpermissions/

**Publisher’s note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s) 2018