LETTER

doi:10.1038/nature13819

A relative shift in cloacal location repositions external genitalia in amniote evolution

Patrick Tschopp[1], Emma Sherratt[2]†, Thomas J. Sanger[2]†, Anna C. Groner[3], Ariel C. Aspiras[1], Jimmy K. Hu[1]†, Olivier Pourquié[1,4,5], Jérôme Gros[6] & Clifford J. Tabin[1]

The move of vertebrates to a terrestrial lifestyle required major adaptations in their locomotory apparatus and reproductive organs. While the fin-to-limb transition has received considerable attention[1,2], little is known about the developmental and evolutionary origins of external genitalia. Similarities in gene expression have been interpreted as a potential evolutionary link between the limb and genitals[3–6]; however, no underlying developmental mechanism has been identified. We re-examined this question using micro-computed tomography, lineage tracing in three amniote clades, and RNA-sequencing-based transcriptional profiling. Here we show that the developmental origin of external genitalia has shifted through evolution, and in some taxa limbs and genitals share a common primordium. In squamates, the genitalia develop directly from the budding hindlimbs, or the remnants thereof, whereas in mice the genital tubercle originates from the ventral and tail bud mesenchyme. The recruitment of different cell populations for genital outgrowth follows a change in the relative position of the cloaca, the genitalia organizing centre. Ectopic grafting of the cloaca demonstrates the conserved ability of different mesenchymal cells to respond to these genitalia-inducing signals. Our results support a limb-like developmental origin of external genitalia as the ancestral condition. Moreover, they suggest that a change in the relative position of the cloacal signalling centre during evolution has led to an altered developmental route for external genitalia in mammals, while preserving parts of the ancestral limb molecular circuitry owing to a common evolutionary origin.

To investigate potential interdependencies between the development of the limbs and external genitalia, we first determined the location of the two structures during embryogenesis. We focused on mouse[4,5] and squamates (lizard and snakes), which show progressive limb reduction[7] yet maintain their external genitalia, the hemipenes[8]. Micro-computed-tomography (μCT) reconstructions of mouse, anole lizard (*Anolis*), python and house snake embryos revealed different anterior–posterior locations of the developing external genitalia relative to limbs. In mice, the genital tubercle is positioned caudal to the hindlimbs (Fig. 1a), whereas in squamates the paired hemipenes bud from the limbs, or from the remnants thereof (Fig. 1b–d). The cloaca, a signalling centre important for genitalia development[6,9,10], is similarly located within the limb-field of squamates (Fig. 1f–h) and expresses *Shh* (Fig. 1j–l). Thus, in squamates all three anatomical structures—limb, hemipenis and cloaca—align at the same anterior–posterior position.

We decided to investigate whether these positional differences would reflect distinct developmental origins of external genitalia. Although the cells of the vertebrate limb bud are known to arise through an epithelial-to-mesenchymal transition (EMT) of an epithelial lateral plate mesoderm (LPM) population lining the coelomic cavity[11], the developmental origin of external genitalia in vertebrates is still unclear. We developed a lentiviral lineage tracing approach (see Methods), to systematically follow the two sources previously proposed, the LPM and tail bud[10,12,13], in three amniote species: mouse, chicken and anole. Injections into the coelom of embryonic day (E)9.5 mouse embryos label cells surrounding the coelom as well as the developing hindlimb (Fig. 2a). However, no green fluorescent protein (GFP)-positive cells are observed in the genital tubercle, with a sharp boundary of labelled cells extending from the coelomic cavity (Fig. 2a, b). In contrast, injection into the posterior mesenchyme of mouse embryos labels the genital tubercle (Fig. 2c). Tail bud injections label the posterior half, whereas the infra-umbilical mesenchyme gives rise to its anterior part[12–14] (Extended Data Fig. 1). In chicken, coelomic injection into stage HH14 embryos labels cells in both limb and genital tubercle (Fig. 2d, e), without any obvious boundary. Tail bud infection also results in GFP-positive genitalia cells, mostly in the posterior tubercle, suggesting that multiple lineages contribute to this species' genitalia (Fig. 2f). Similar conclusions were reached in a parallel study[31]. In *Anolis*, coelomic injections at stage 2–3 result in GFP-positive cells in the limb and the developing genitalia (Fig. 2g, h), whereas no labelled cells are seen in the hemipenes after tail bud injections (Fig. 2i).

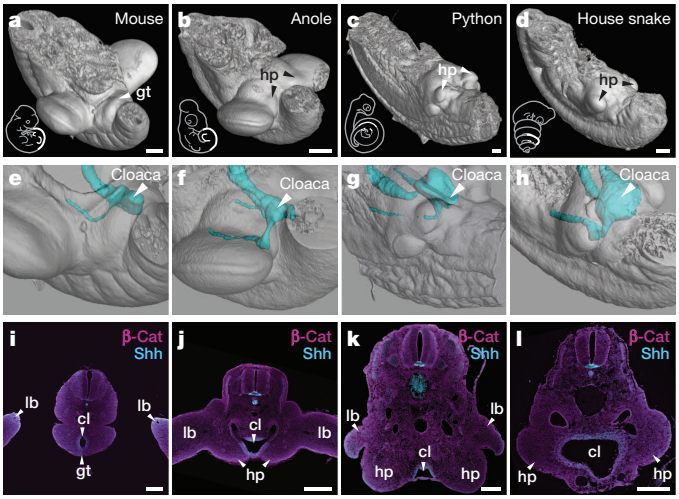


Figure 1 | A relative positional shift of limbs, genitalia and the cloaca in squamates. a–d, μCT scans of mouse (**a**), anole (**b**), python (**c**) and house snake (**d**) lumbosacral regions (highlighted in sketch in white) at embryonic stages, illustrating the position of the developing external genitalia. **e–h**, Three-dimensional reconstructions of cloacal volumes. The cloaca is located at the same anterior–posterior position as the limb in squamates (**f–h**); however, it is positioned more posteriorly in the mouse (**e**). **i–l**, Transversal sections stained for β-catenin and Shh, indicating the conservation of a cloacal signalling centre in all four species. gt, genital tubercle; hp, hemipenis; lb, limb; cl, cloaca. Scale bars, 200 μm.

[1]Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA. [2]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138, USA. [3]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA. [4]Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), 67400 Illkirch, France. [5]Department of Pathology, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA. [6]Developmental and Stem Cell Biology Department, Institut Pasteur, 75724 Paris Cedex 15, France. †Present addresses: Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, Iowa 50011, USA (E.S.); Department of Molecular Genetics and Microbiology, University of Florida, Gainesville, Florida 32610, USA (T.J.S.); Department of Orofacial Sciences and Program in Craniofacial and Mesenchymal Biology, UCSF, San Francisco, California 94143, USA (J.K.H.).

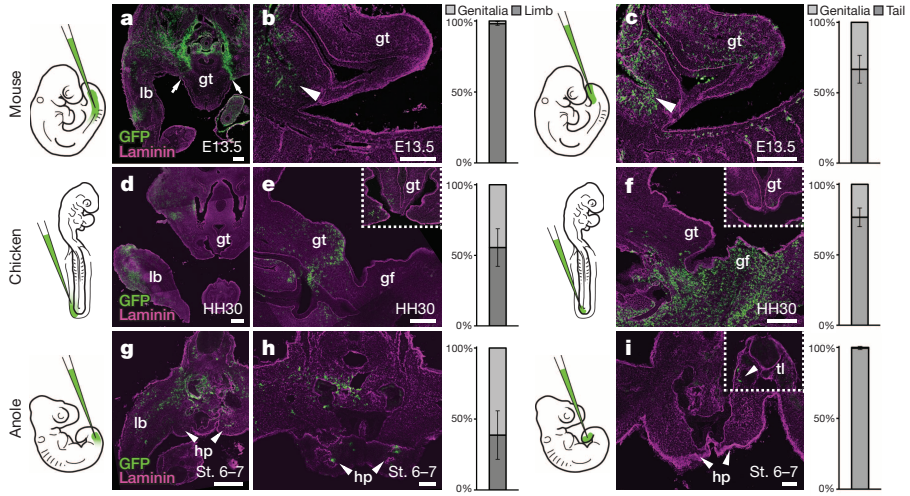


Figure 2 | Differential developmental origins of external genitalia in amniotes. **a**–**i**, Transversal and sagittal views of GFP lentivirus-injected embryos. Relative contribution of GFP-positive cells to respective organs is quantified on the right, normalized on tissue area. Error bars represent standard deviation in at least $n = 4$ biological replicates. **a**, **b**, Injection into the coelom at mouse E9.5 ($n = 48$) labels the limb at E13.5, but excludes the genital tubercle (arrows). Only cells lining the peritoneal cavity are labelled (**b**, arrowhead), but none in the genital tubercle proper. **c**, Injection into the tail bud ($n = 101$) labels cells in the genital tubercle. Accidental piercing of the coelom labels cells of the peritoneal cavity (arrowhead). **d**, **e**, Coelom injection in HH14 chicken embryos ($n = 81$) labels the limb and the genital tubercle at HH30. **e**, Sagittal and transversal close-up (inset) views. **f**, Sagittal and transversal close-up (inset) views of tail-bud-injected chick embryos ($n = 77$), showing labelling in the genital tubercle. **g**, **h**, Anole embryos injected into the coelom at stage (St.) 2–3 ($n = 94$) show GFP labelling of both limb and hemipenis at stage 6–7. **i**, No hemipenis cells are labelled following tail bud injection ($n = 57$), even though there are GFP-positive cells in the tail (inset, arrowhead). gt, genital tubercle; gf, genital fold; hp, hemipenis; lb, limb; tl, tail. Scale bars, 200 μm (**a**–**g**), 50 μm (**h**, **i**).

As for chicken and mouse limbs[11], cells of the *Anolis* limb, but also the hemipenis, originate through an EMT of the coelomic epithelium (Extended Data Fig. 2a, b). In snakes, we find evidence for similar cellular dynamics. The hemipenes emerge as small buds at a 'limb-like' lateral position, juxtaposed to the coelomic cavity (Extended Data Fig. 2c–f). A concomitant basement membrane breakdown, consistent with an EMT of the LPM, is seen in the budding of both mouse limbs and snake hemipenes (Extended Data Fig. 2d, f). Moreover, we find that *Tbx4*, a gene important for hindlimb development[15] is expressed from early on, in both the coelomic epithelium and the mesenchyme of the developing hemipenis (Extended Data Fig. 2g–i). Its forelimb counterpart *Tbx5* is expressed later, in the mesenchyme only, in agreement with its pattern of expression in the genitalia of mammals[16] (Extended Data Fig. 2j–l). This suggests that squamate external genitalia initiate with limb-like cellular dynamics, with the resulting mesenchymal cell population in modern snakes being converted to a genital fate[17]. Thus, important differences exist in the developmental origins of external genitalia in amniotes: chicken genitals originate from both LPM and the tail bud, whereas the mouse genital tubercle consists of infra-umbilical mesenchyme and tail-bud-descendant cells. In contrast, the hemipenis shares a developmental route with hindlimbs, either through secondary budding as in lizards, or by entirely recruiting the mesenchymal cell population to a genital fate in modern snakes. Given the impact of developmental lineage on an organ's molecular architecture, we next explored the transcriptomes of emerging genitalia, in the two opposing trajectories of mammals and squamates.

Gene expression profiling has successfully been applied to questions of developmental and evolutionary origin, of cell types and entire morphological features[18–21]. We thus dissected the early and late stages of developing limbs and genitalia from mouse and anole embryos (Fig. 3a) for comparative RNA-sequencing (RNA-seq) analyses. Overall transcriptome similarities were assessed using multidimensional scaling (MDS; Fig. 3b). The transcriptomes dominantly resolve along dimension 1 in a species-dependent manner, as expected for similar tissues in evolutionarily distant species[22]. Dimension 2, however, contains a clear organ identity signal, that is, limb versus genitalia. This separation is virtually absent in *Anolis* samples compared with mouse, reflecting similar transcriptional programs owing to a common developmental origin for genitalia and hindlimbs. For hierarchical clustering, we included the early tail bud as outgroup of the primary body axis, and the forelimbs to account for anterior–posterior differences in the two species' genitalia (Fig. 3c, d; see Methods). In *Anolis*, the early hemipenis transcriptome falls within the limb clade, indicative of an almost generic limb molecular architecture, and only later differentiates into a more organ-specific signature (Fig. 3c and Extended Data Fig. 3a). In contrast, mouse genitalia transcriptomes are clearly distinct from limbs, from early on, highlighting the separate developmental origins of the two organs (Fig. 3d and Extended Data Fig. 3b).

To identify genes driving hindlimb- and genitalia-specific transcriptome separation, we used principal component analysis. Principal component 1 (PC1) correlates with species differences, whereas organ specificity is resolved along PC2 (Fig. 3e), allowing us to identify organ-specific 'driver' genes in a species-independent manner. We assessed the contribution of orthologous genes to PC2, according to their absolute loading values (Fig. 3f). Gene ontology (GO) analysis[23] of the top 500 genes (Fig. 3f and Supplementary Table 1) identified GO terms related to transcription factors and signalling molecules (Fig. 3g). Gene regulatory networks thus determine limb versus genital organ transcriptomes, but also mirror their developmental origin. Notably, transcription factor/signalling molecule data are sufficient to reproduce the clustering seen with whole transcriptomes (Fig. 3h, i and Extended Data Fig. 4a, b). Collectively, we find a clear distinction between mouse genitalia and limb transcriptomes during early and late organogenesis. Such genitalia-specific separation is only seen in late *Anolis* hemipenes, arguing for a developmental repurposing to a copulatory structure. Importantly, we find transcriptional similarities between early hemipenes and hindlimbs in squamates, illustrating a common developmental origin.

An attractive model for the varied developmental origins of amniote external genitalia would be a repositioning of the cloacal signalling centre with respect to different mesenchymal cell populations with progenitor potential. Hence, bringing either hindlimb or tail bud close to the cloaca would allow these lineages to contribute to genital outgrowth. We tested this hypothesis by grafting GFP-transgenic chicken or quail cloacae into the hindlimb bud of wild-type embryos (Fig. 4a and Extended Data

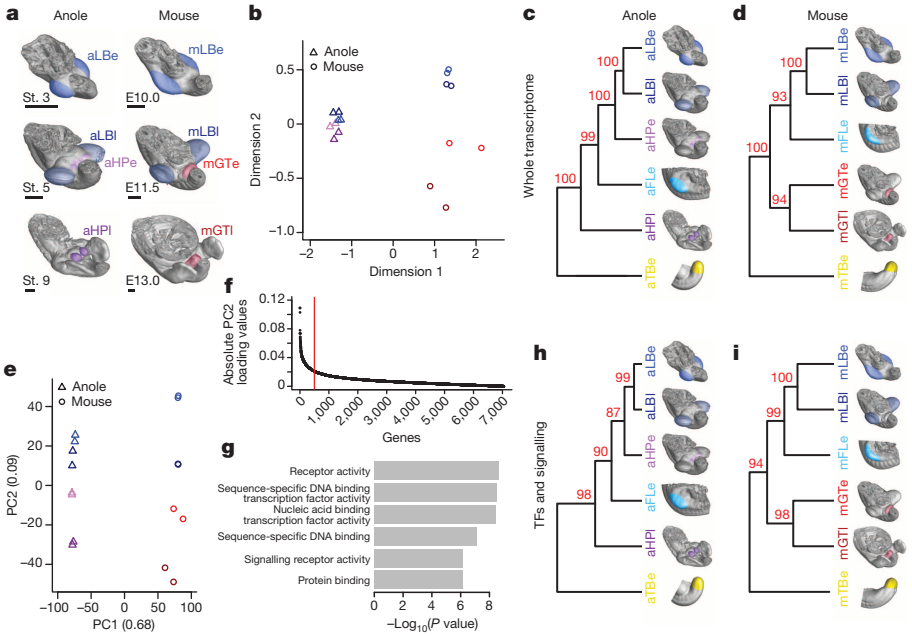


Figure 3 | Molecular architecture of limbs and external genitalia in lizards and mice. **a**, Micro-dissected tissues for RNA-seq analysis, highlighted by colour code and four-letter sample identifier. Early and late limb buds and genitalia buds were analysed in anole lizard and mouse embryos ($n = 2$). **b**, MDS analysis reveals greater overall transcriptome similarities in anole limb and genitalia data sets (triangles) than in their mouse counterparts (circles). **c**, **d**, Hierarchical clustering of pairwise Pearson's correlation coefficients for whole-transcriptome data from anole (**c**) and mouse (**d**) samples. Additional data sets are stage (St.) 2–3 anole and E9.5 mouse forelimb (turquoise) and tail bud (yellow). Numbers at nodes represent approximately unbiased *P* values, obtained by multiscale bootstrap resampling. **e**, Principal component analysis. Species transcriptomes separate along PC1, whereas organs are resolved along PC2. Numbers in brackets indicate proportion of variance, explained by PC1 and PC2. **f**, Absolute loading values for PC2, as shown in **e**. **g**, GO term enrichment analysis using the top 500 genes (red line in **f**). Top hits include transcription-factor- and signalling-pathway-related terms. **h**, **i**, Hierarchical clustering analysis of pairwise Pearson's correlation coefficients for transcription factors (TFs) and signalling pathways data from anole (**h**) and mouse (**i**) samples. Sample identifiers: a, anole; m, mouse; GT, genital tubercle; HP, hemipenis; LB, limb; e, early; l, late.

Fig. 5a–c). After 1–2 days of incubation, limbs showed ectopic, secondary buds (Fig. 4b, c). GFP-negative cells suggest an inductive effect of the cloaca on the surrounding mesenchyme, rather than simple overproliferation of the graft itself (Fig. 4c). Similar buds could be induced by grafting beads soaked in the known cloacal signalling molecules SHH and FGF[9,10,24,25] (Extended Data Fig. 5d–g). To assess the fate of responding cells, we re-analysed our RNA-seq data for potential genital versus limb markers. For both species, we performed stage- and organ-specific differential expression analyses. Of the 2,003 genes showing an absolute $\log_2$ fold change greater than 1.5 (P value < 0.05) (Extended Data Fig. 6), we identified 27 that are altered in all four comparisons, 25 of which were altered in the same direction (Fig. 4d). Hierarchical clustering of normalized mouse expression values reveals four stage- and organ-specific signatures, which are largely conserved in *Anolis* (Fig. 4e). On the basis of expression patterns and levels (Extended Data Figs 6 and 7), we chose marker genes to assess transcriptional changes due to ectopic cloacal signals. Indeed, chicken limb cells close to GFP-positive cloacal grafts downregulate limb markers *LHX9* and *TBX18* (Fig. 4f, g), and ectopically express genital markers *ISL1* (ref. 26), *GATA2* and *RUNX1*

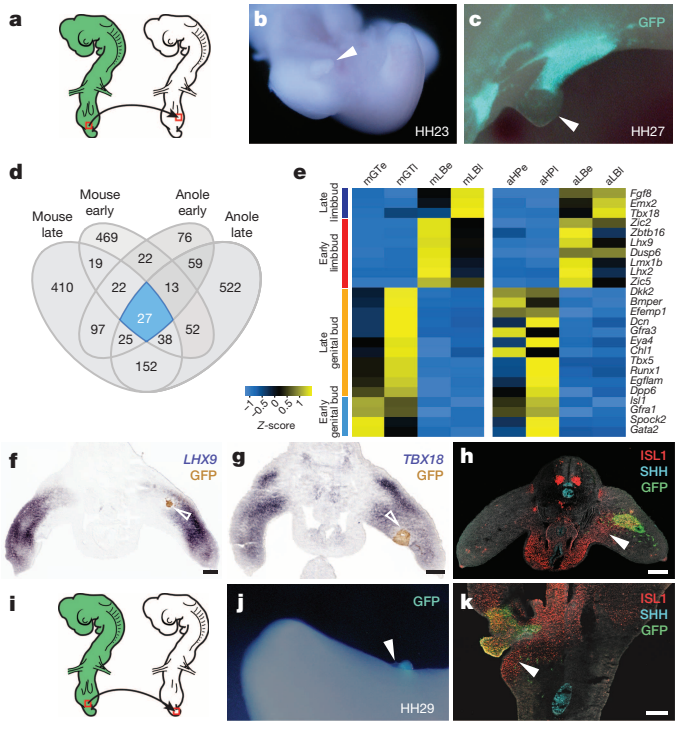


Figure 4 | The cloacal signalling centre can recruit different mesenchymal cell populations for the outgrowth of external genitalia. **a**, Schematic of the hindlimb grafting procedure in chicken embryos. GFP-transgenic cloacae are transplanted into the proximal-ventral portion of wild-type hindlimbs. **b**, **c**, Ectopic outgrowth (arrowheads) on limbs with cloacal grafts ($n = 30/118$). **d**, Venn diagram of pairwise differential expression analysis results (log(fold change) > 1.5, P value < 0.05) of limbs versus genital tissues for early and late budding stages in anole and mouse. **e**, Heat map of Z-score-normalized values of core 25 genes showing consistent differential expression between limbs and genitalia. Mouse row-based hierarchical clustering was re-used for anole samples. **f**–**h**, Analysis of limb- and genital-specific markers. Expression of limb markers *LHX9* (**f**) and *TBX18* (**g**) is downregulated near the GFP-positive cloacal graft (open arrowheads), while genital marker ISL1 (**h**) is expressed ectopically (arrowhead). **i**, Schematic of the tail-bud-grafting procedure in chicken embryos. GFP-transgenic cloacae are transplanted into ventral wild-type tail buds. **j**, Ectopic outgrowth on tails with cloacal grafts ($n = 16/87$). **k**, Genital marker ISL1 is upregulated ectopically in the tail, close to the GFP-positive cloacal graft (arrowhead). All gene expression was assessed in at least $n = 3$ samples. Scale bars, 200 μm.

(Fig. 4h and Extended Data Fig. 5h–k). Accordingly, when we graft cloacal tissue into the tail bud mesenchyme, ectopic budding and genital marker expression are equally induced (Fig. 4i–k and Extended Data Figs 5l–n and 8). Hence, mesenchymal cells from different developmental origins can respond to inductive cloacal signals, generating outgrowths and genitalia-like marker gene expression. Importantly, these results support the idea that changing the relative anterior–posterior position of the cloaca could generate external genitalia with distinct developmental origins during the course of amniote evolution.

In summary, we show substantial variation in external genitalia development in extant amniote species. We propose that repositioning the cloaca can recruit different mesenchymal cell populations, either through spatial or heterochronic changes in the dynamics of their emergence. In squamates, the hindlimb is the dominant source, with modern snakes entirely repurposing a mesenchymal bud to a genital fate. In mice, limbs and genitals have discrete developmental origins—the LPM and the ventral and tail bud mesenchyme, respectively—with chicken showing an intermediate state. Moreover, we find that similarities in limb and genitalia transcriptomes are dependent on the cellular source of the primordia from which they emerge. Specifically, there is a higher degree of early transcriptome congruence in species deriving their intromittent organs from limb anlagen. Notably, the ability of different mesenchymal cell populations to respond to cloacal, genitalia-inducing signals seems conserved in extant species. It is therefore tempting to speculate that a limb-derived state could represent the ancestral condition in the evolution of external genitalia, as suggested by their position relative to limbs during turtle development[17,27] and the bifid genitalia of basal mammals[28,29]. As such, a developmental continuity between limbs and genitalia could have turned into an 'evolutionary continuity' in mammals, as the two organs spatially separated owing to a relative repositioning of the cloaca[30]. Once-shared developmental trajectories could thus help to explain molecular similarities still noticeable in species that now develop the limb and genitalia from distinct cellular sources[3,4,19].

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

Received 10 March; accepted 20 August 2014.
Published online XX 2014.

1. Shubin, N., Tabin, C. & Carroll, S. Fossils, genes and the evolution of animal limbs. *Nature* **388,** 639–648 (1997).
2. Wagner, G. P. & Chiu, C. H. The tetrapod limb: a hypothesis on its origin. *J. Exp. Zool.* **291,** 226–240 (2001).
3. Kondo, T., Zakany, J., Innis, J. W. & Duboule, D. Of fingers, toes and penises. *Nature* **390,** 29 (1997).
4. Yamada, G. *et al.* Molecular genetic cascades for external genitalia formation: an emerging organogenesis program. *Dev. Dyn.* **235,** 1738–1752 (2006).
5. Cohn, M. J. Development of the external genitalia: conserved and divergent mechanisms of appendage patterning. *Dev. Dyn.* **240,** 1108–1115 (2011).
6. Lin, C. *et al.* Delineating a conserved genetic cassette promoting outgrowth of body appendages. *PLoS Genet.* **9,** e1003231 (2013).
7. Greer, A. E. Limb reduction in squamates: identification of the lineages and discussion of the trends. *J. Herpetol.* **25,** 166–173 (1991).
8. Cope, E. D. On the hemipenes of the Sauria. *Proc. Acad. Nat. Sci. Philadelphia* **48,** 461–467 (1896).
9. Haraguchi, R. *et al.* Unique functions of sonic hedgehog signaling during external genitalia development. *Development* **128,** 4241–4250 (2001).
10. Perriton, C. L., Powles, N., Chiang, C., Maconochie, M. K. & Cohn, M. J. Sonic hedgehog signaling from the urethral epithelium controls external genital development. *Dev. Biol.* **247,** 26–46 (2002).
11. Gros, J. & Tabin, C. J. Vertebrate limb bud formation is initiated by localized epithelial-to-mesenchymal transition. *Science* **343,** 1253–1256 (2014).
12. Ohta, S., Suzuki, K., Tachibana, K., Tanaka, H. & Yamada, G. Cessation of gastrulation is mediated by suppression of epithelial-mesenchymal transition at the ventral ectodermal ridge. *Development* **134,** 4315–4324 (2007).
13. Suzuki, K., Economides, A., Yanagita, M., Graf, D. & Yamada, G. New horizons at the caudal embryos: coordinated urogenital/reproductive organ formation by growth factor signaling. *Curr. Opin. Genet. Dev.* **19,** 491–496 (2009).
14. Matsumaru, D. *et al.* Genetic analysis of the role of *Alx4* in the coordination of lower body and external genitalia formation. *Eur. J. Hum. Genet.* **22,** 350–357 (2014).
15. Naiche, L. A. Loss of Tbx4 blocks hindlimb development and affects vascularization and fusion of the allantois. *Development* **130,** 2681–2693 (2003).
16. Chapman, D. L., Garvey, N., Hancock, S. & Alexiou, M. Expression of the T-box family genes, *Tbx1–Tbx5*, during early mouse development. *Dev. Dyn.* **206,** 379–390 (1996).
17. Raynaud, A. & Pieau, C. in *Biology of the Reptilia, Development B* (eds Gans, C. & Billett, F. S.) Vol. 15, 149–300 (John Wiley & Sons, 1985).
18. Arendt, D. Genes and homology in nervous system evolution: comparing gene functions, expression patterns, and cell type molecular fingerprints. *Theory Biosci.* **124,** 185–197 (2005).
19. Wagner, G. P. The developmental genetics of homology. *Nature Rev. Genet.* **8,** 473–479 (2007).
20. Shubin, N., Tabin, C. & Carroll, S. Deep homology and the origins of evolutionary novelty. *Nature* **457,** 818–823 (2009).
21. Wang, Z., Young, R. L., Xue, H. & Wagner, G. P. Transcriptomic analysis of avian digits reveals conserved and derived digit identities in birds. *Nature* **477,** 583–586 (2011).
22. Merkin, J., Russell, C., Chen, P. & Burge, C. B. Evolutionary dynamics of gene and isoform regulation in mammalian tissues. *Science* **338,** 1593–1599 (2012).
23. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. *Nature Genet.* **25,** 25–29 (2000).
24. Haraguchi, R. *et al.* Molecular analysis of external genitalia formation: the role of fibroblast growth factor (*Fgf*) genes during genital tubercle formation. *Development* **127,** 2471–2479 (2000).
25. Miyagawa, S. *et al.* Dosage-dependent hedgehog signals integrated with Wnt/β-catenin signaling regulate external genitalia formation as an appendicular program. *Development* **136,** 3969–3978 (2009).
26. Suzuki, K. *et al.* Reduced BMP signaling results in hindlimb fusion with lethal pelvic/urogenital organ aplasia: a new mouse model of sirenomelia. *PLoS ONE* **7,** e43453 (2012).
27. Larkins, C. E. & Cohn, M. J. Phallus development in the turtle *Trachemys scripta*. *Sex Dev.* http://dx.doi.org/10.1159/000363631 (2014).
28. Griffiths, M. in *The Biology of the Monotremes* (Academic, 1978).
29. De Barros, M. A. *et al.* Marsupial morphology of reproduction: South America opossum male model. *Microsc. Res. Tech.* **76,** 388–397 (2013).
30. Jurberg, A. D., Aires, R., Varela-Lasheras, I., Novoa, A. & Mallo, M. Switching axial progenitors from producing trunk to tail tissues in vertebrate embryos. *Dev. Cell* **25,** 451–462 (2013).
31. Herrera, A. M. & Cohn, M. J. Embryonic origin and compartmental organization of the external genitalia. *Sci. Rep.* (in the press).

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

Acknowledgements The authors thank D. Duboule, H. Kaessmann, A. Necsulea, B. Okaty, G. Rey and G. P. Wagner for discussions, M. A. de Bakker for the snake *Tbx5* probe and A. M. Herrera and M. J. Cohn for discussing and sharing unpublished results. μCT scans were performed at the Center for Nanoscale Systems, Harvard University (supported by National Science Foundation award ECS-0335765) and at the Museum of Comparative Zoology. Next-generation sequencing was performed at the HMS Biopolymers Facility and computational analyses were run on the Orchestra Cluster, HMS Research Computing. P.T. was supported by post-doctoral fellowships from the Swiss National Science Foundation, EMBO and the Human Frontiers Science Program. A.C.G. was supported by a post-doctoral fellowship from the Swiss National Science Foundation. This work was supported by National Institutes of Health grant R37-HD032443 to C.J.T.

Author Contributions P.T., J.G. and C.J.T. conceived the project and designed the experiments. P.T. performed most experiments and computational analyses. E.S. prepared CT scans and helped with statistical analyses. T.J.S. helped with CT scans, *Anolis* husbandry and embryo collection. A.C.G. produced lentiviruses and A.C.A. helped with grafting experiments. J.K.H., O.P. and J.G. initiated snake analyses. O.P. contributed snake embryos. J.G. contributed to chick lineage tracing experiments. P.T., J.G. and C.J.T. wrote the paper, with comments from co-authors.

Author Information Sequencing data has been deposited in the Gene Expression Omnibus under accession number GSE60373. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.J.T. (tabin@genetics.med.harvard.edu) or J.G. (jgros@pasteur.fr).

METHODS

Tissue sample collection. All embryos were collected in accordance with the appropriate Institutional Animal Care and Use Committee (IACUC) guidelines. Timed-pregnant CD1 and C57BL/6 females were purchased from Charles River Laboratories. Gravid *Anolis* females were purchased from Candy Quality Reptiles (IACUC #26-11 and #28-14). Anole housing and egg incubation were done as previously described[32]. To collect early stage pre-oviposition embryos, females were euthanized by intraperitoneal Euthasol injection and eggs were dissected from the oviduct. Snake husbandry and egg collection have been described before[33]. Fertilized white leghorn chicken eggs were obtained from Charles River Laboratories and incubated at 38 °C. For staging of mouse embryos, noon on the day of the vaginal plug was considered as E0.5. For chicken and anole embryos, staging was performed according to Hamburger and Hamilton[34] or Sanger and colleagues[32], respectively. Embryonic tissue was dissected in cold PBS and either fixed in 4% PFA and processed for cryo-embedding or CT scanning, or else directly processed for RNA extraction or stored in RNAlater (Qiagen).

CT scanning and image processing. Embryos were fixed in 4% PFA and stored in 100% ethanol. Staining was done for 48 h in 30% PTA (1% (w/v) phosphotungstic acid in water) and 70% ethanol[35]. Specimens were rinsed and stored in 70% ethanol until image acquisition. CT scans were carried out on a Bruker Skyscan 1173 or a Nikon (Metris) X-tek HMXST225, at 50–57 kV, 115–145 μA and 810–1,000 ms exposure time. Voxel sizes ranged from 0.0023 to 0.0056, with 1,500 to 2,400 total projections. Post-processing of scan data was done in VGStudio MAX 2.2 (Volume Graphics). For three-dimensional reconstruction of the cloaca, serial TIFF stacks produced by VGStudio were read into the Imaris software package (Bitplane) and the endodermal epithelium was used as a guide to manually outline the extent of the cloacal volume.

Immunohistochemistry and *in situ* hybridization. Fixed embryos were embedded in 7.5% gelatin/15% sucrose or dehydrated in sucrose gradients and embedded in OCT. Sectioning was performed on a Leica CM3000 cryostat. For immunohistochemistry, sections were incubated with primary antibodies in PBST (PBS/BSA 0.2%, Triton 0.1%, SDS 0.02%) overnight, washed twice for 10 min in PBST and incubated for 1 h with secondary antibodies. To detect genital tubercle and ectopic limb Isl1 expression, as well as Lmx1b, the signal was amplified using the TSA Plus Cy3 kit (Perkin Elmer). Primary antibodies used were anti-β-catenin (BD Biosciences), anti-Shh (Santa Cruz), anti-laminin (Sigma), anti-GFP (Abcam), anti-QCPN, anti-Lmx1b and anti-Isl1 (all Developmental Studies Hybridoma Bank). *In situ* hybridization was performed using standard protocols[36]. Fluorescent images were acquired on a Zeiss LSM10 inverted confocal miscroscope. Bright-field images were acquired on a Nikon Eclipse E1000. Whole-mount images were acquired on a Leica MZ FL III. Images were globally processed for colour balance and brightness using Adobe Photoshop.

Lineage tracing analyses. For lentiviral lineage tracing, viral particles harbouring ubiquitously expressing GFP cassettes (UbiC-GFP or hPGK-GFP) were produced by transient transfection in 293T cells as described elsewhere[37]. Viral particles were then injected into either the coelomic cavitiy, or the tail bud mesenchyme of mouse E9.5 embryos, chicken HH14 or *Anolis* stage 2–3. For mouse experiments, timed-pregnant CD1 females were anaesthetized using isoflurane and surgery, *in utero* visualization of embryos and virus injection was done as previously outlined[38]. For chicken embryos, eggs were lowered and windowed and virus injected using a pressure injector. For *Anolis* lineage tracing, we developed a novel whole-embryo *ex ovo* culturing system, using media conditions previously described for squamate organ cultures[39]. Briefly, we prepared culture dishes with an indentation by pouring 1% Agar Noble (BD Difco) dissolved in culture medium into cell culture dishes containing a modelling clay, egg-shaped casting mould. Once solidified, the mould was removed and stage 2–3 *Anolis* embryos, dissected in 2× PBS, were placed with their yolk intact in the resulting cavity and covered with culture medium. 10% ink in 1× PBS/PenStrep was mouth-pipetted underneath the embryo, for better visualization, and lentiviral particles were injected using a pressure injector. To increase viral infection rate, embryo plates were kept for 12–16 h at 37 °C in a humidified chamber, before switching them to 28 °C. Additional tail bud injections were performed using DiI. Embryos survived for up to 12 days. Only live specimens showing overall normal morphology were considered for further analysis.

To assess GFP^+ cell contribution, embryos were dissected, fixed in 4% PFA, gelatin-embedded and cryo-sectioned. Sections were stained for GFP and imaged on a Zeiss LSM10 inverted confocal miscroscope. For quantifications, 4–5 embryos per condition and species were imaged on multiple sections spanning the respective organs. GFP^+ cell counting was performed in ImageJ, using the 'ITCN' plugin written by T. Kuo (UCSB). A total of 59,331 GFP^+ cells were counted (mouse: 33,853; chicken: 23,710; anole: 1,768). Counts were averaged over multiple sections and normalized on tissue area measured. The resulting ratios of the two tissues are given as a percentage of total GFP^+ cells per area.

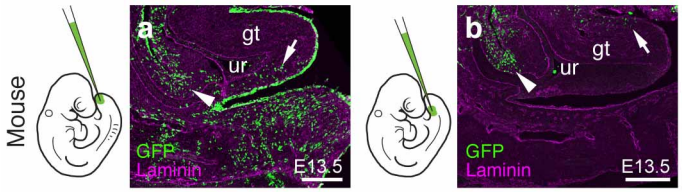
RNA sample preparation and sequencing. Total RNA was extracted from freshly dissected or RNAlater-preserved tissue using the Arcturus PicoPure RNA Isolation Kit (Life Technologies) and enriched for mRNA fraction using MPG mRNA Purification Kit (PureBiotech). Multiplexed RNA-seq libraries were produced using the SPIA cDNA synthesis kit and the Ovation Ultralow DR Multiplex system (NuGen). Sequencing was done on an Illumina HiSeq 2000, with eight samples multiplexed per lane. Base calling was performed using the Illumina software. Over 561,000,000 50-bp reads were generated, with an average of 23.4 million reads per sample (median = 22,000,000 reads). For all tissues, biological duplicates were sequenced.

Read mapping and transcriptome analyses. Initial read mapping was performed with the RNA-seq unified mapper (RUM)[40], using mouse NCBI37/mm9 and *Anolis* AnoCar2.0 genome assemblies, with UCSC mm9 refseq and ASU_Acar v2.1 (ref. 41) annotation files, respectively. This resulted in 10,265 orthologous genes between mouse and *Anolis*. An in-house improved annotation for *Anolis*, generated before the publication of Eckalbar *et al.*[41], yielded concordant results in all downstream analyses. To account for species-specific differences in non-uniquely mapping (NU) reads, we redistributed NU reads based on the number of uniquely mapping (U) reads mapping to the respective loci, following a logic outlined before[42]. Multi-dimensional scaling analysis was carried out on normalized read counts using the edgeR bioconductor package[43]. Genes differentially expressed between early fore- and hindlimb samples were determined in edgeR, and genes showing consistent changes in mouse and *Anolis* were excluded from further analyses, to dampen potential anterior–posterior differences between the organs. For hierarchical clustering, correlation coefficient and principal component analyses, we calculated transcripts per million (TPM)[44] values, which were then $\log_2$-transformed. Hierarchical clustering was done using the 'pvclust' R package[45], with 1,000 iterations of multi-scale bootstrap re-sampling, and approximately unbiased (AU) *P* values are provided in the graph. Heat maps of correlation coefficients were plotted with the 'lattice' R package. Principal component analysis was done with the 'prcomp' function in the 'stats' R package and 'GOseq'[46] was used for GO-term enrichment analysis. Pairwise differential expression analysis between early and late budding stages, in mouse and *Anolis* limbs and genitalia, was done in edgeR, and Venn diagram of genes with an absolute log(fold change) > 1.5 and P value < 0.05 was visualized using VennDiagram[47].

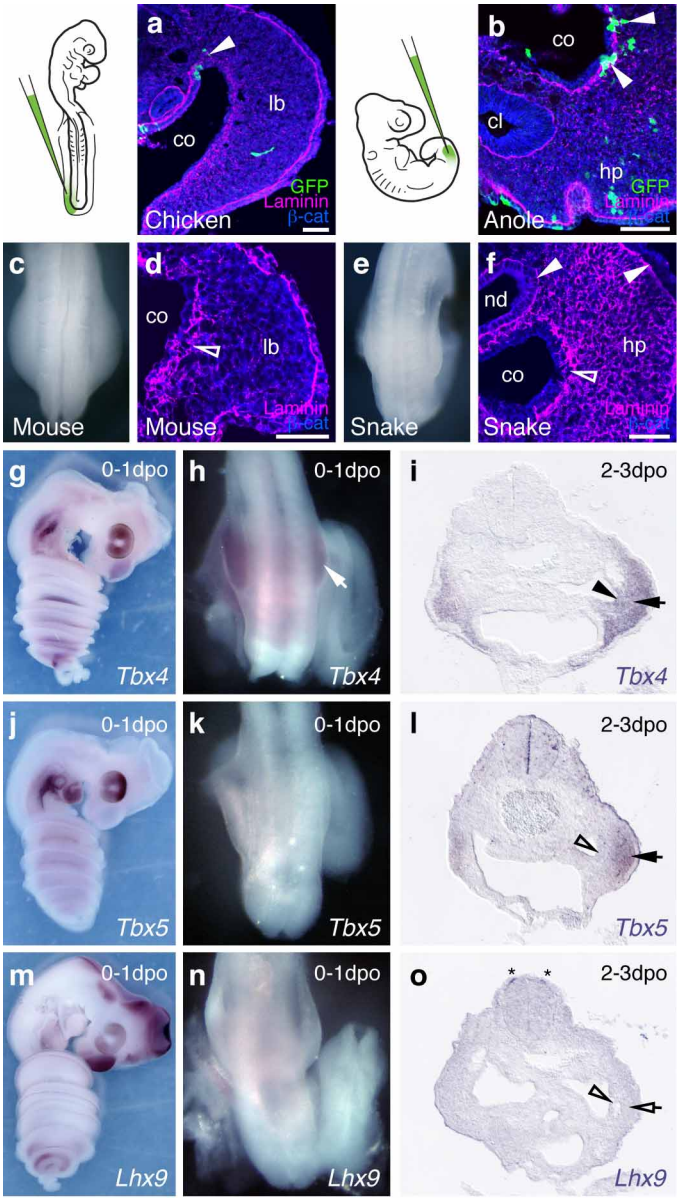
Grafting experiments. For heteroptopic, homochronic cloacal grafts, donor and recipient embryos were incubated to reach stage HH17–20. Donor embryos were either GFP-transgenic chicken[48], purchased from Clemson University, or quail, purchased from Strickland GameBird Farm. Cloacas were dissected in ice-cold PBS and grafted using tungsten needles to a proximal-ventral position, to mimic the squamate configuration, and removed from the apical ectodermal ridge to avoid *SHH*-induced digit duplications[10], or the tail bud in wild-type recipient chicken. Successful grafts were incubated for 1–3 additional days, dissected and screened for the appearance of ectopic outgrowths. Donor versus recipient tissue was discriminated using either GFP or QCPN antibody staining on cryo-sections, or GFP fluorescence for whole-mount embryos. Sham surgery or grafting of GFP-positive limb mesenchyme did not cause any comparable outgrowths. Cloaca-induced outgrowths never stained positive for Alcian blue at later stages, indicating that they were not digit duplications (data not shown). For bead experiments, Affi-Gel Blue Gel beads (150–300 μm; Bio-Rad) were washed in PBS and incubated for 1–2 h at room temperature, in PBS with recombinant proteins (SHH, FGF2, FGF8; all R&D Systems) at concentrations of 0.1–1 μg μl^{-1}. Soaked beads were briefly washed in PBS and grafted to limb and tail buds, as outlined for the cloacal grafts. Control grafts using beads soaked in PBS with bovine serum albumin did not yield any observable outgrowths.

32. Sanger, T. J., Losos, J. B. & Gibson-Brown, J. J. A developmental staging series for the lizard genus *Anolis*: a new system for the integration of evolution, development, and ecology. *J. Morphol.* **269,** 129–137 (2008).
33. Gomez, C. *et al.* Control of segment number in vertebrate embryos. *Nature* **454,** 335–339 (2008).
34. Hamburger, V. & Hamilton, H. L. A series of normal stages in the development of the chick embryo. *J. Morphol.* **88,** 49–92 (1951).
35. Metscher, B. D. MicroCT for developmental biology: a versatile tool for high-contrast 3D imaging at histological resolutions. *Dev. Dyn.* **238,** 632–640 (2009).
36. McGlinn, E. & Mansfield, J. H. Detection of gene expression in mouse embryos and tissue sections. *Methods Mol. Biol.* **770,** 259–292 (2011).
37. Barde, I., Salmon, P. & Trono, D. *Production and Titration of Lentiviral Vectors* (John Wiley & Sons, 2001).
38. Punzo, C. & Cepko, C. L. Ultrasound-guided *in utero* injections allow studies of the development and function of the eye. *Dev. Dyn.* **237,** 1034–1042 (2008).
39. Buchtová, M. *et al.* Initiation and patterning of the snake dentition are dependent on Sonic Hedgehog signaling. *Dev. Biol.* **319,** 132–145 (2008).
40. Grant, G. R. *et al.* Comparative analysis of RNA-Seq alignment algorithms and the RNA-Seq Unified Mapper (RUM). *Bioinformatics* **27,** 2518–2528 (2011).
41. Eckalbar, W. L. *et al.* Genome reannotation of the lizard *Anolis carolinensis* based on 14 adult and embryonic deep transcriptomes. *BMC Genomics* **14,** 49 (2013).

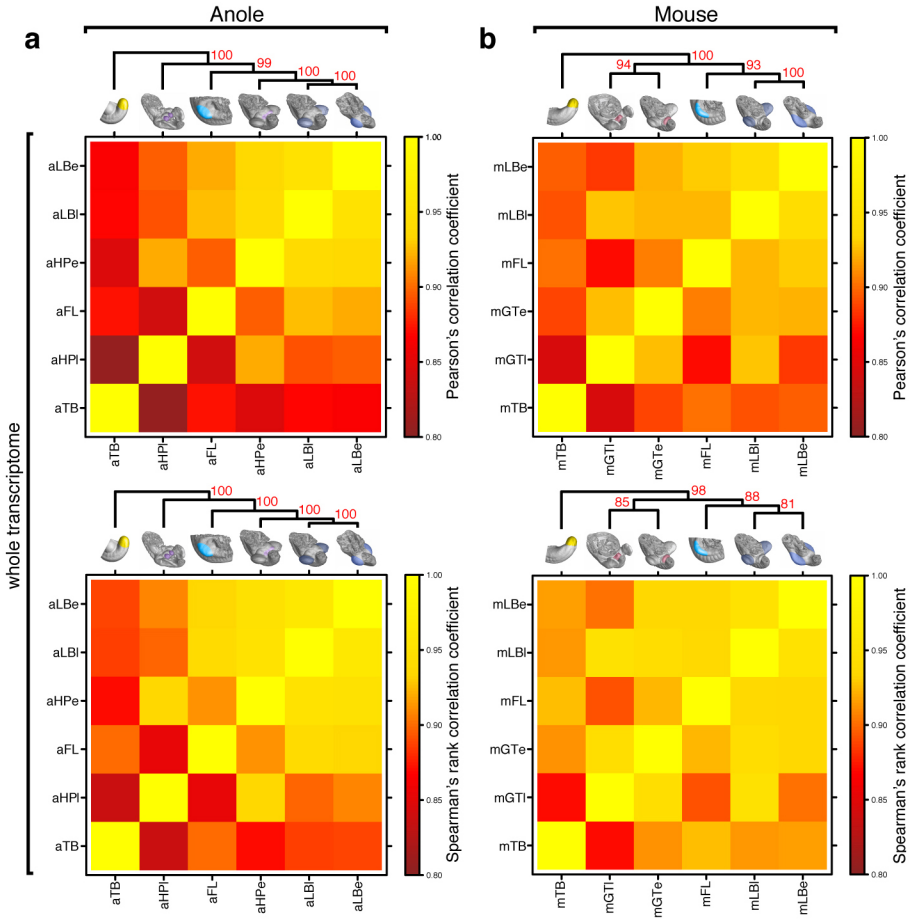
42. Brawand, D. *et al.* The evolution of gene expression levels in mammalian organs. *Nature* **478,** 343–348 (2011).
43. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26,** 139–140 (2010).
44. Wagner, G. P., Koryu, K. & Lynch, V. J. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory Biosci.* **131,** 281–285 (2012).
45. Suzuki, R. & Shimodaira, H. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* **22,** 1540–1542 (2006).
46. Young, M. D., Wakefield, M. J., Smyth, G. K. & Oshlack, A. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* **11,** R14 (2010).
47. Chen, H. & Boutros, P. C. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics* **12,** 35 (2011).
48. Chapman, S. C. *et al.* Ubiquitous GFP expression in transgenic chickens using a lentiviral vector. *Development* **132,** 935–940 (2005).



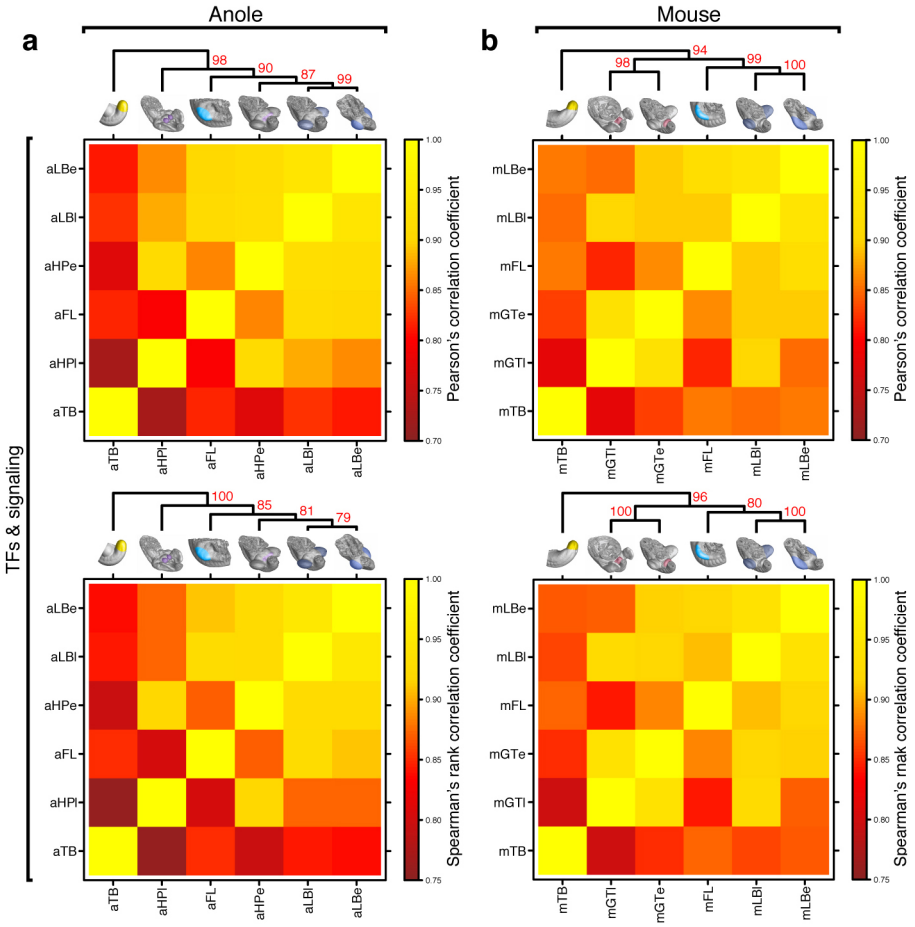
Extended Data Figure 1 | Two separable ventral cell populations give rise to the murine genital tubercle. **a**, **b**, Injection into the most distal ventral part of the embryo, the tail bud, marks cells posterior/ventral to the phallic part of the urethra (**a**, arrow; $n = 7$), whereas injection closer to the allantois, into the infra-umbilical mesenchyme, labels cells anterior/dorsal to the phallic part of the urethra (**b**, arrow; $n = 4$). Cells lining the peritoneal cavity are also marked (arrowheads), owing to accidental piercing of the coelom. gt, genital tubercle; ur, urethra. Scale bars, 200 μm.



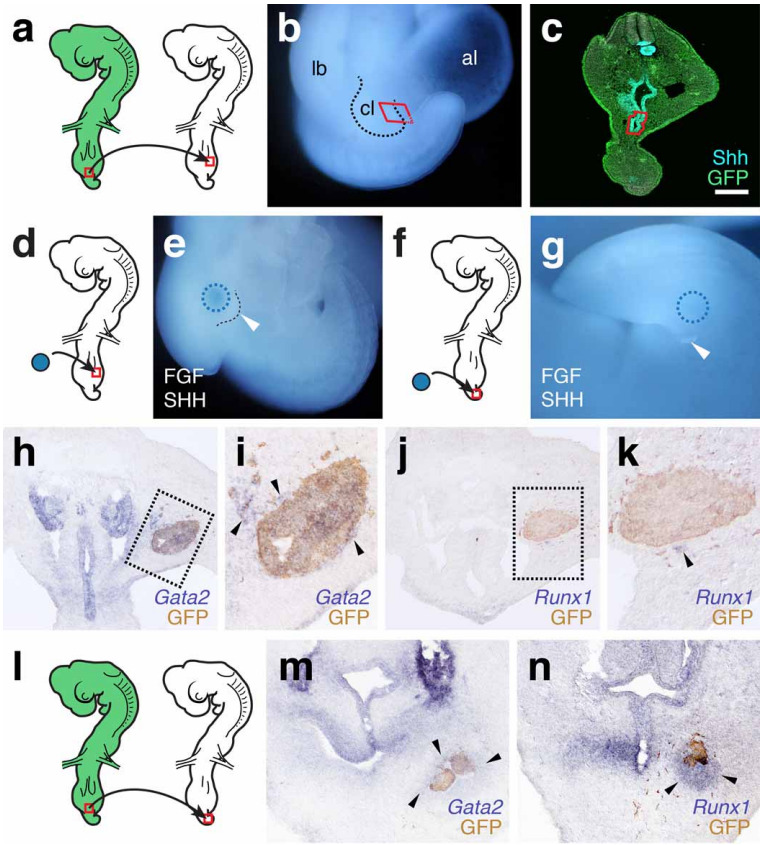
Extended Data Figure 2 | The squamate hemipenis mesenchyme initiates with limb-like cellular dynamics from the coelomic epithelium through an EMT. **a**, Injection of GFP-expressing lentiviruses into the coelom of chicken embryos at HH14 labels cells emerging from the epithelium that contribute to the hindlimb mesenchyme (arrowhead). **b**, In lizards, labelled cells leaving the coelomic epithelium contribute to the hemipenis mesenchyme (arrowheads). **c**, Dorsal view of the hindlimb region of an E10.0 mouse embryo. **d**, Transversal section of a limb bud, showing EMT of the coelomic epithelium (diffuse laminin staining, open arrowhead), as cells contribute to the limb-bud mesenchyme. **e**, Dorsal view of the budding hemipenis of a snake embryo, 1 day after egg deposition. **f**, Transversal section of the hemipenis region. The basement membrane of the coelomic epithelium is breaking down (open arrowhead), while it is intact for both the nephric duct and the surface ectoderm (arrowheads). **g**–**o**, Expression of genitalia and limb genes during hemipenis initation. **g**–**i**, *Tbx4* is expressed early (**h**, arrow) and late during hemipenis initiation, in both the coelomic epithelium (**i**, arrowhead) and the hemipenis mesenchyme (**i**, arrow). **j**–**l**, *Tbx5* is only expressed later, in the mesenchyme (**l**, arrow), but is absent from the coelomic epithelium (**l**, open arrowhead). **m**–**o**, Limb marker gene *Lhx9* (see also Fig. 4e) is absent from both epithelium (**o**, open arrowhead) and mesenchyme (**o**, open arrow), but can be detected in dI1 neurons (**o**, asterisk). All gene expression was assessed in at least $n = 3$ samples. cl, cloaca; co, coelom; hp, hemipenis; lb, limb; nd, nephric duct. Scale bars, 50 µm.



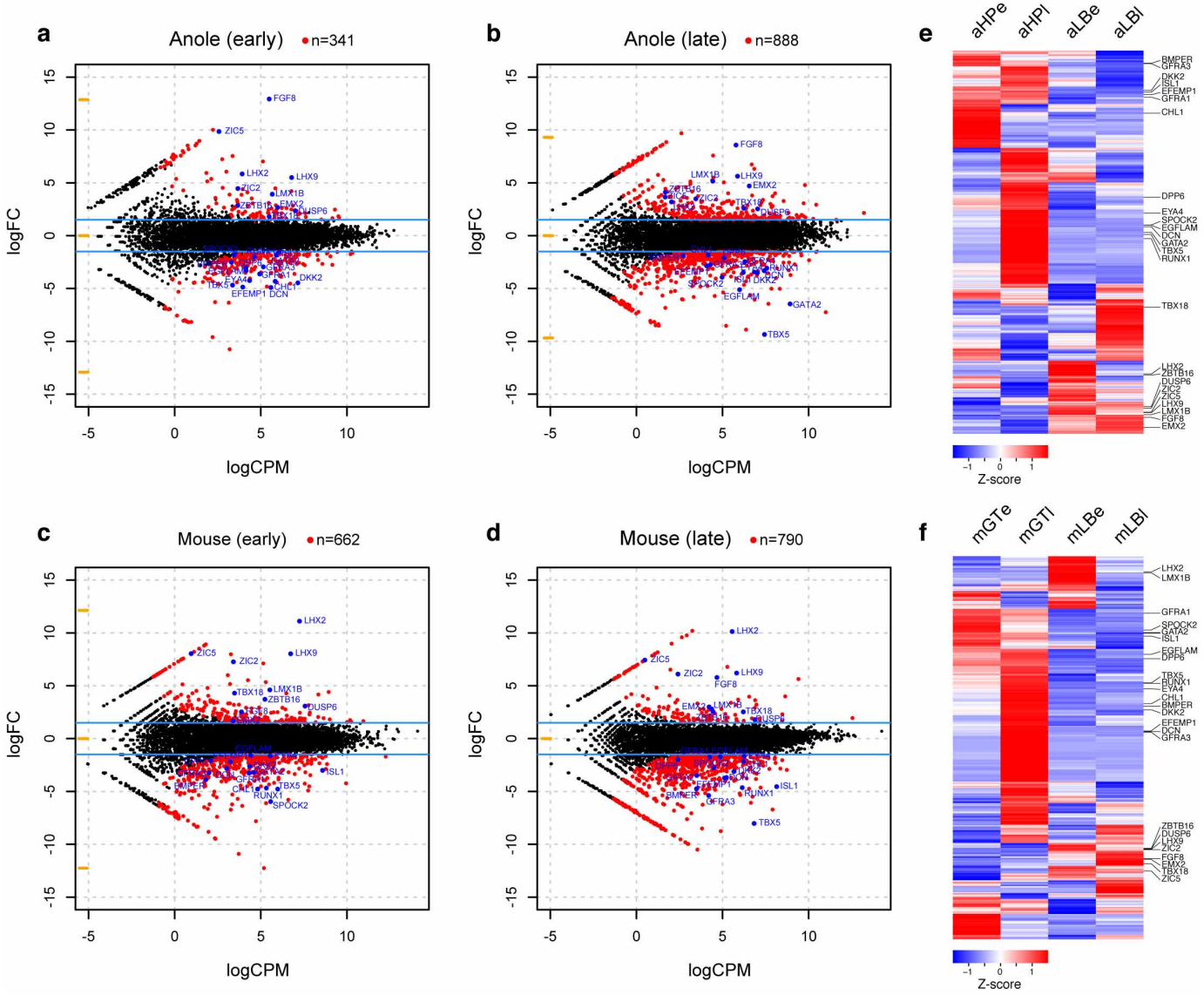
Extended Data Figure 3 | Heat maps of Pearson's and Spearman's rank correlation coefficients and cluster analysis of whole-transcriptome data. **a**, **b**, Hierarchical clustering on pairwise correlation coefficients for whole-transcriptome data from anole (**a**) and mouse (**b**) samples. Numbers at nodes represent approximately unbiased *P* values obtained by multiscale bootstrap resampling. Sample identifiers: a, anole; m, mouse; GT, genital tubercle; HP, hemipenis; LB, limb; e, early; l, late.



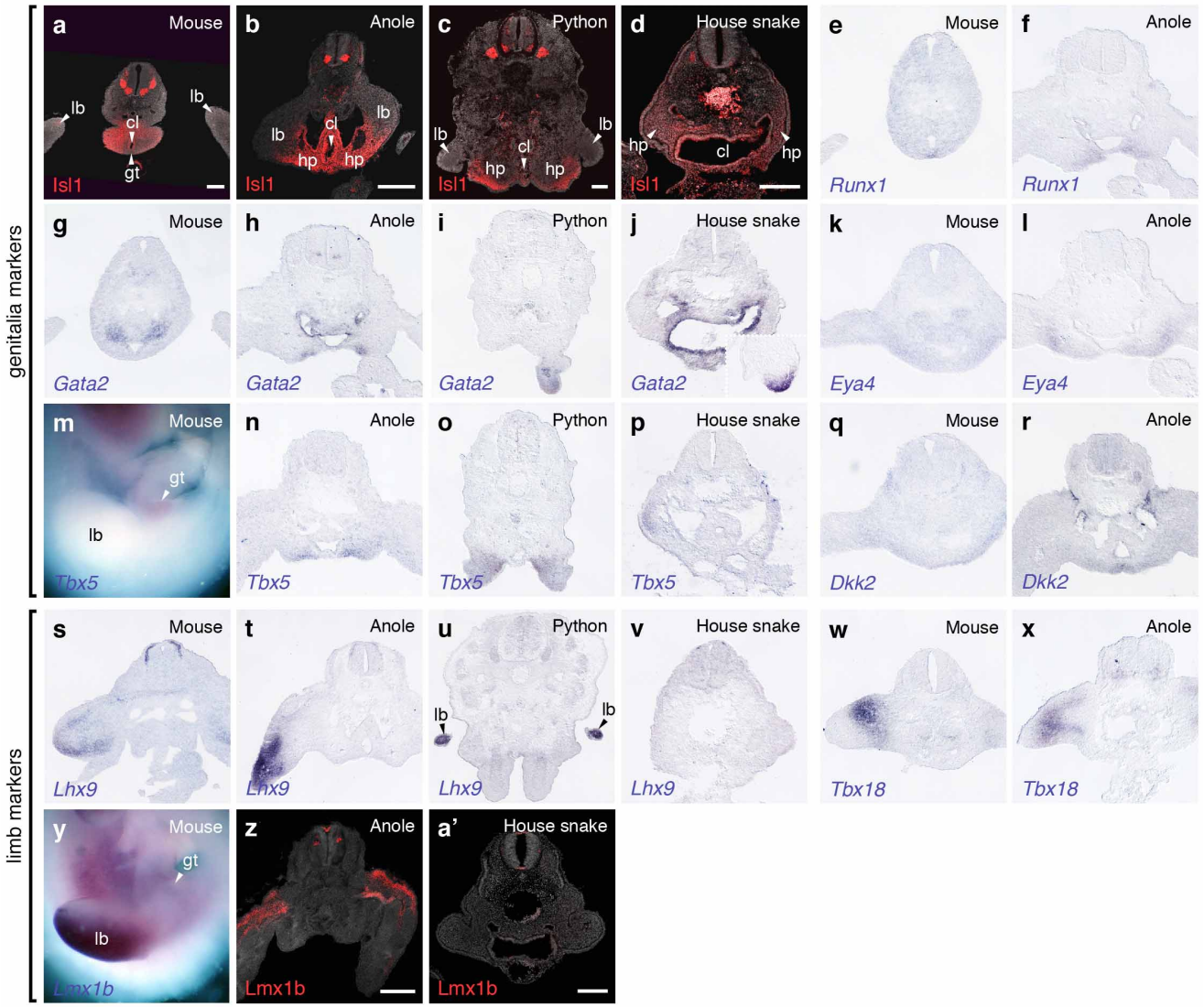
Extended Data Figure 4 | Heat maps of Pearson's and Spearman's rank correlation coefficients and cluster analysis of transcription factor and signalling pathway data. **a**, **b**, Hierarchical clustering on pairwise correlation coefficients of transcription factor (TF) and signalling pathway data from anole (**a**) and mouse (**b**) samples. Numbers at nodes represent approximately unbiased *P* values obtained by multiscale bootstrap resampling. Sample identifiers: a, anole; m, mouse; GT, genital tubercle; HP, hemipenis; LB, limb; e, early; l, late.



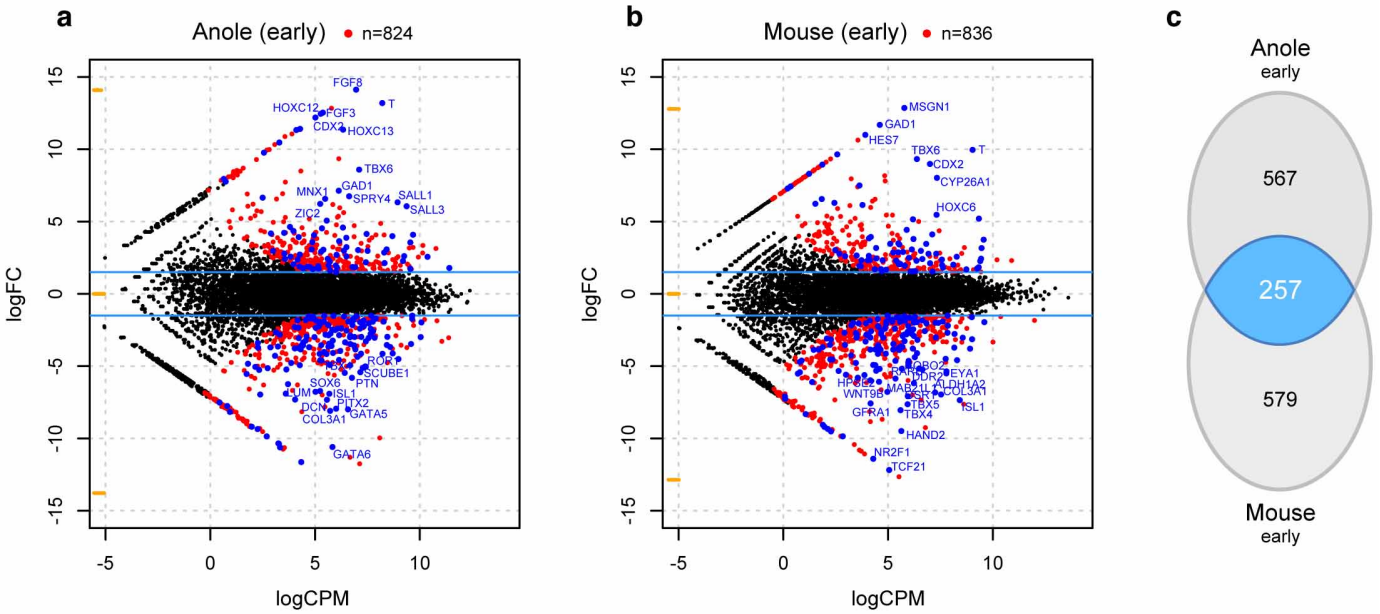
Extended Data Figure 5 | Heterotopic grafting of the cloacal signalling centre leads to ectopic outgrowths and genitalia-like transcriptional changes. **a**–**c**, Schematics and close-up images of the cloacal grafting procedure. **a**, The cloaca of a stage HH17–19 GFP-transgenic chicken embryo (red rectangle) is transplanted into the proximal-ventral portion of the limb of a wild-type embryo. **b**, **c**, Only the ventral-most part of the cloaca, including the cloacal membrane, is dissected out (**b**, red box), and subsequently cleared of excess mesenchymal cells attached to the *SHH*-expressing endoderm (**c**, red outline). **d**–**g**, Grafting of beads soaked in SHH and FGF can induce ectopic outgrowths on both limbs (**e**; $n = 6/48$) and tail (**g**; $n = 3/31$). **h**–**k**, Ectopic expression of genital markers *GATA2* (**h**, **i**, arrowheads) and *RUNX1* (**j**, **k**, arrowhead) in limb buds, following cloaca-to-limb grafts. **l**–**n**, Ectopic expression of genital marker *GATA2* (**m**, arrowheads) and *RUNX1* (**n**, arrowheads) in the tail region, following cloaca-to-tail grafts. All gene expression was assessed in at least $n = 3$ samples. al, allantois; cl, cloaca; lb, limb. Scale bar, 200 μm.



Extended Data Figure 6 | Pairwise differential expression analysis of limb and genitalia transcriptomes. **a**–**d**, Smear plot visualization of differential expression analyses of early anole (**a**), late anole (**b**), early mouse (**c**) and late mouse (**d**) limb versus genitalia transcriptomes. Genes used for the Venn diagram in Fig. 4d ($|\log_2(\text{fold change})| > 1.5$; P value < 0.05) are highlighted in red, core 25 marker genes (see Fig. 4e and text) are highlighted and labelled in blue. CPM, counts per million; FC, fold change. **e**, **f**, Heat map of Z-score-normalized expression values for all genes fulfilling Venn diagram criteria ($n = 2{,}003$), for anole (**e**) and mouse (**f**) data. Row-based hierarchical clustering was used; core 25 marker genes are indicated on the right.



Extended Data Figure 7 | Comparative marker gene expression analysis in mouse and squamate embryos. **a**–**r**, Genitalia markers *Isl1* (**a**–**d**), *Runx1* (**e**, **f**), *Gata2* (**g**–**j**), *Eya4* (**k**, **l**), *Tbx5* (**m**–**p**) and *Dkk2* (**q**, **r**). *Gata2* only becomes visibly expressed at the later stages of house snake hemipenis development (**j**, inset). **s**–**a**′, Limb markers *Lhx9* (**s**–**v**), *Tbx18* (**w**, **x**) and *Lmx1b* (**y**–**a**′). All gene expression was assessed in at least $n = 3$ samples. cl: cloaca; gt: genital tubercle; hp: hemipenis; lb: limb. Scale bar, 200 μm.



Extended Data Figure 8 | Pairwise differential expression analysis of tail bud and genitalia transcriptomes. **a**, **b**, Smear plot visualization of differential expression analyses of early anole (**a**) and early mouse (**b**) tail bud versus genitalia transcriptomes. Genes used as input for the Venn diagram in **c** ($|\log_2(\text{fold change})| > 1.5$; P value < 0.05) are highlighted in red, overlapping 257 marker genes are highlighted in blue. Top 25 genes in the two species, based on logCPM (counts per million) and logFC (fold change), are labelled. **c**, Venn diagram showing overlap of pairwise differential expression analysis results (log(fold change) > 1.5, P value < 0.05) of tail bud versus genital tissues for early budding stages in both anole and mouse.