

impossible. It is likely that all traits influenced and provided the context for the evolution of others (31).

## REFERENCES AND NOTES

- J. Gauthier, L. F. Gall, Eds., *New Perspectives on the Origin and Early Evolution of Birds* (Peabody Museum of Natural History, New Haven, CT, 2001).
- G. Dyke, G. Kaiser, Eds., *Living Dinosaurs: the Evolutionary History of Modern Birds* (Wiley, Chichester, UK, 2011).
- J. Gauthier, *Mem. Calif. Acad. Sci.* **8**, 1–55 (1986).
- P. C. Sereno, *Annu. Rev. Earth Planet. Sci.* **25**, 435–489 (1997).
- K. Padian, A. J. de Ricqlès, J. R. Horner, *Nature* **412**, 405–408 (2001).
- T. A. Dececchi, H. C. E. Larsson, *Evolution* **67**, 2741–2752 (2013).
- F. E. Novas, M. D. Ezcurra, F. L. Agnolín, D. Pol, R. Ortiz, *Rev. Mus. Argentino de Cienc. Nat. n.s.* **14**, 57–81 (2012).
- A. H. Turner, D. Pol, J. A. Clarke, G. M. Erickson, M. A. Norell, *Science* **317**, 1378–1381 (2007).
- M. T. Carrano, in *Arniote Paleobiology*, M. T. Carrano, T. J. Gaudin, R. W. Blob, J. R. Wible, Eds. (Univ. of Chicago Press, Chicago, 2006), chap. 8.
- D. W. E. Hone, T. M. Keeseey, D. Pisani, A. Purvis, *J. Evol. Biol.* **18**, 587–595 (2005).
- R. B. Sookias, R. J. Butler, R. B. J. Benson, *Proc. Biol. Sci.* **279**, 2180–2187 (2012).
- B.-A. S. Bhullar et al., *Nature* **487**, 223–226 (2012).
- A. M. Heers, K. P. Dial, *Trends Ecol. Evol.* **27**, 296–305 (2012).
- V. Allen, K. T. Bates, Z. Li, J. R. Hutchinson, *Nature* **497**, 104–107 (2013).
- M. N. Puttick, G. H. Thomas, M. J. Benton, *Evolution* **68**, 1497–1510 (2014).
- R. B. J. Benson et al., *PLOS Biol.* **12**, e1001853 (2014).
- D. K. Zelenitsky et al., *Science* **338**, 510–514 (2012).
- X. Zheng et al., *Science* **339**, 1309–1312 (2013).
- R. B. J. Benson, J. N. Choiniere, *Proc. Biol. Sci.* **280**, 20131780 (2013).
- Materials and methods are available as supplementary materials on Science Online. Data files are archived on Dryad Digital Repository (doi:10.5061/dryad.jm6pj).
- P. Godefroit et al., *Nature* **498**, 359–362 (2013).
- X. Xu et al., *Nature* **484**, 92–95 (2012).
- J. N. Choiniere et al., *Science* **327**, 571–574 (2010).
- P. Christiansen, R. A. Farina, *Hist. Biol.* **16**, 85 (2004).
- A. J. Drummond, M. A. Suchard, D. Xie, A. Rambaut, *Mol. Biol. Evol.* **29**, 1969–1973 (2012).
- A. J. Drummond, S. Y. W. Ho, M. J. Phillips, A. Rambaut, *PLOS Biol.* **4**, e88 (2006).
- M. Pagel, A. Meade, *BayesTraits: Software and Documentation* (2013); www.evolution.reading.ac.uk/BayesTraitsV2Beta.html.
- L. E. Zanno, P. J. Makovicky, *Proc. Biol. Sci.* **280**, 20122526 (2013).
- J. Hanken, D. B. Wake, *Annu. Rev. Ecol. Syst.* **24**, 501–519 (1993).
- O. W. M. Rauhut, C. Foth, H. Tischlinger, M. A. Norell, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 11746–11751 (2012).
- T. S. Kemp, *Proc. Biol. Sci.* **274**, 1667–1673 (2007).

## ACKNOWLEDGMENTS

We thank e-research SA for use of high-performance computing facilities, the Environment Institute (University of Adelaide) and Australian Research Council for funding, and the University of Bologna for logistic support. Author contributions: A.C., G.J.D., and D.N. collected phylogenetic and stratigraphic data; M.S.Y.L. analyzed data and assisted with data collection; all authors contributed to interpretation of results; M.S.Y.L. wrote and revised manuscript with input from all other authors. Data files are archived on Dryad Digital Repository (doi:10.5061/dryad.jm6pj).

## SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6196/562/suppl/DC1  
Materials and Methods  
Figs. S1 to S9  
References (32–62)

14 February 2014; accepted 20 June 2014  
10.1126/science.1252243

## MODELING DIGITS

# Digit patterning is controlled by a Bmp-Sox9-Wnt Turing network modulated by morphogen gradients

J. Raspopovic,<sup>1\*</sup> L. Marcon,<sup>1\*</sup> L. Russo,<sup>1</sup> J. Sharpe<sup>1,2†</sup>

During limb development, digits emerge from the undifferentiated mesenchymal tissue that constitutes the limb bud. It has been proposed that this process is controlled by a self-organizing Turing mechanism, whereby diffusible molecules interact to produce a periodic pattern of digital and interdigital fates. However, the identities of the molecules remain unknown. By combining experiments and modeling, we reveal evidence that a Turing network implemented by Bmp, Sox9, and Wnt drives digit specification. We develop a realistic two-dimensional simulation of digit patterning and show that this network, when modulated by morphogen gradients, recapitulates the expression patterns of Sox9 in the wild type and in perturbation experiments. Our systems biology approach reveals how a combination of growth, morphogen gradients, and a self-organizing Turing network can achieve robust and reproducible pattern formation.

Digits form in a periodic pattern that alternates digital and interdigital fates along the anterior-posterior (AP) axis of the limb bud. Traditionally, this pattern has been explained by a positional information model (1) based on an AP gradient of Sonic hedgehog (Shh) (2, 3). However, embryonic and genetic manipulations (4, 5) have shown that digit patterning is independent of Shh and may be instead controlled by a self-organizing mechanism. Over three decades ago, it was proposed that such a mechanism could be a Turing system (6, 7), in which a diffusible activator and inhibitor (8) interact and self-organize to form the periodic digit pattern. Recent work has strengthened this hypothesis (9); however, two important questions remain to be addressed.

First, although a number of mathematical Turing models have been proposed to explain the periodic digit pattern (10, 11), no computer simulation has been able to correctly reproduce the expression patterns of digit markers over time and space. Second, the diffusible molecules that implement the Turing network have not yet been identified. Transforming growth factor- $\beta$  (TGF- $\beta$ ) molecules were proposed as activators in a Turing system (12, 13), but the corresponding diffusible inhibitor could not be found. The galectins CG-1A and CG-8 have also been proposed as Turing molecules in the chick (14) but are not believed to play a similar role in mammals (15). More recently, bone morphogenetic proteins (BMPs) and their receptors have been proposed to form a Turing network (10), but this model assumed that the BMP receptors can diffuse through tissue, for which there is no evidence.

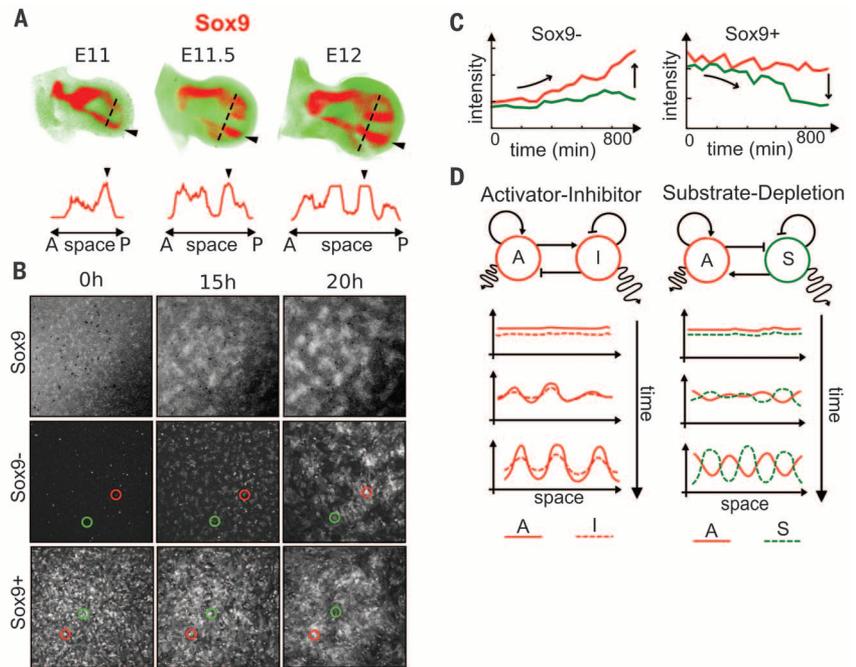
To identify the molecules that control digit specification, it is crucial to distinguish the genes involved in early digit patterning from downstream differentiation factors. We therefore analyzed the expression of the earliest known skeletal marker *Sox9* (16) and identified embryonic day 11.5 (E11.5) as the earliest stage that shows a periodic digital pattern (Fig. 1A). We also performed micromass cultures with E11.5 Sox9-EGFP (enhanced green fluorescent protein) limb autopods (17) and found that cells create a periodic Sox9 pattern by 15 hours, faster than previously reported (18), with dynamics reminiscent of a two-dimensional (2D) Turing simulation (Fig. 1B and movie S1). We observed that the periodic pattern formed even when the culture was initiated with Sox9<sup>-</sup> or Sox9<sup>+</sup> cells sorted by fluorescence-activated cell sorting (FACS) (Fig. 1, B and C, and movies S2 and S3), confirming that Sox9 is dynamically regulated by a self-organizing patterning mechanism.

When *Sox9* is knocked out, all the genes that reflect a digital or interdigital pattern (e.g., *Bmp2*, *Chordin*, *Noggin*) lose their normal periodic expression (19), suggesting that *Sox9* itself is part of the Turing network rather than a downstream differentiation marker. The two simplest Gierer-Meinhardt topologies (8), the activator-inhibitor and the substrate-depletion model (Fig. 1D), predict that the diffusible molecules of the Turing network should have periodic patterns that are either in-phase or out-of-phase, respectively. We therefore performed a microarray analysis to identify genes related to the major developmental signaling pathways that were differentially expressed between Sox9<sup>+</sup> and Sox9<sup>-</sup> FACS-sorted cells. On the basis of both the number of genes and fold-change magnitude, the pathways that were most strongly represented were WNT, BMP, and FGF (fibroblast growth factor) (fig. S1). We performed a second level of screening by whole-mount in situ hybridization (WMISH) to check which genes were genuinely expressed in-phase

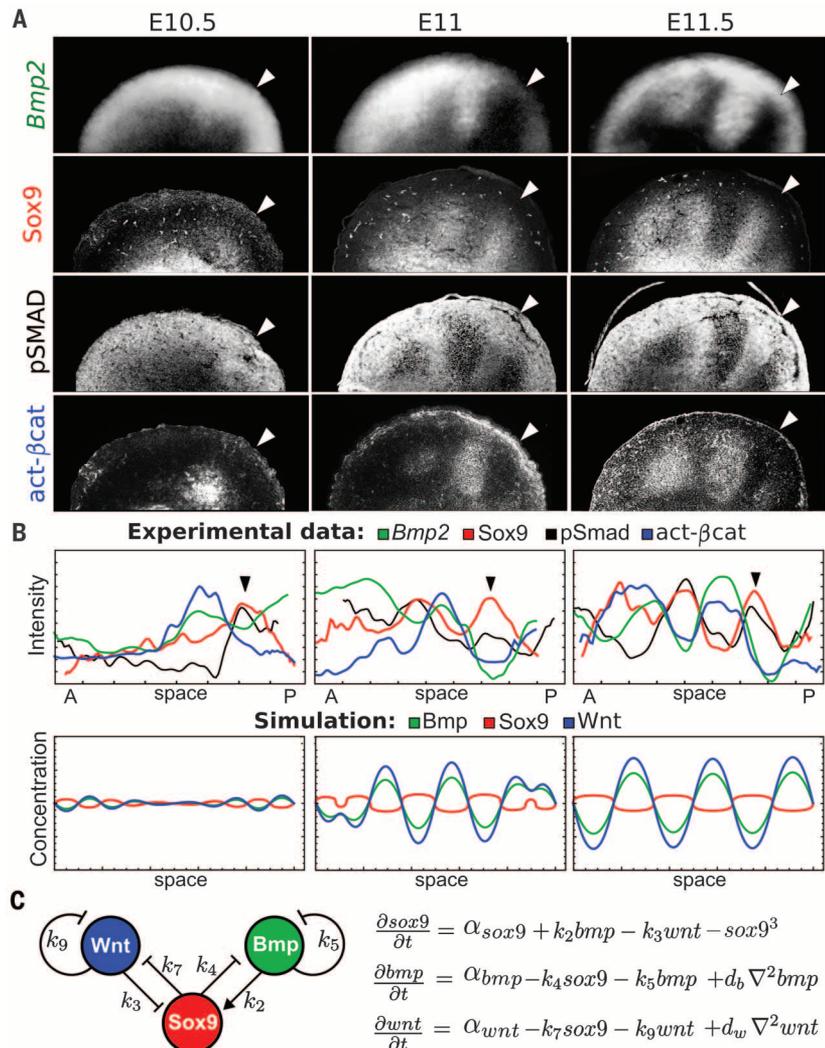
<sup>1</sup>Systems Biology Program, Centre for Genomic Regulation (CRG), and Universitat Pompeu Fabra (UPF), Dr. Aiguader 88, 08003 Barcelona, Spain. <sup>2</sup>Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys 23, 08010 Barcelona, Spain.

\*These authors contributed equally to this work. †Corresponding author. E-mail: james.sharpe@crgeu

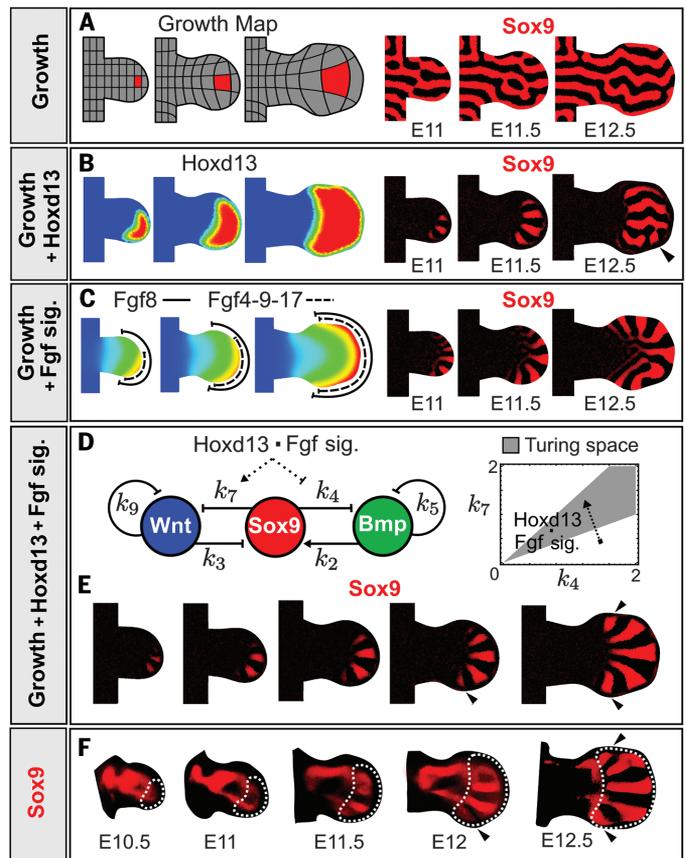
**Fig. 1. The Sox9 periodic digit pattern.** (A) Top: Sox9 expression patterns (red) from E11 to E12 (limb shape in green). Bottom: Anterior-posterior (AP) profiles (dashed line above) highlight the progressive appearance of the Sox9 periodic pattern, first visible at E11.5; arrowheads indicate digit 4. Limb orientation in all figures: distal to the right, posterior to the bottom. (B) First row: time course of micromass culture with Sox9-EGFP autopod cells; a periodic pattern emerges by 15 hours. Second and third rows: When cultures are initiated with Sox9<sup>-</sup> or Sox9<sup>+</sup> cells obtained by FACS sorting, they still form a periodic pattern by up-regulating or down-regulating Sox9 depending on the position (single cells highlighted by red and green circles). (C) Sox9-EGFP intensities of the tracked cells over time. (D) The two simplest models that can form a Turing pattern. In the Activator-Inhibitor model, an autocatalytic activator (A) promotes its own inhibitor (I), resulting in in-phase periodic patterns. In the Substrate-Depletion model, an activator (A) activates itself by depleting its own substrate (S), resulting in out-of-phase periodic patterns.



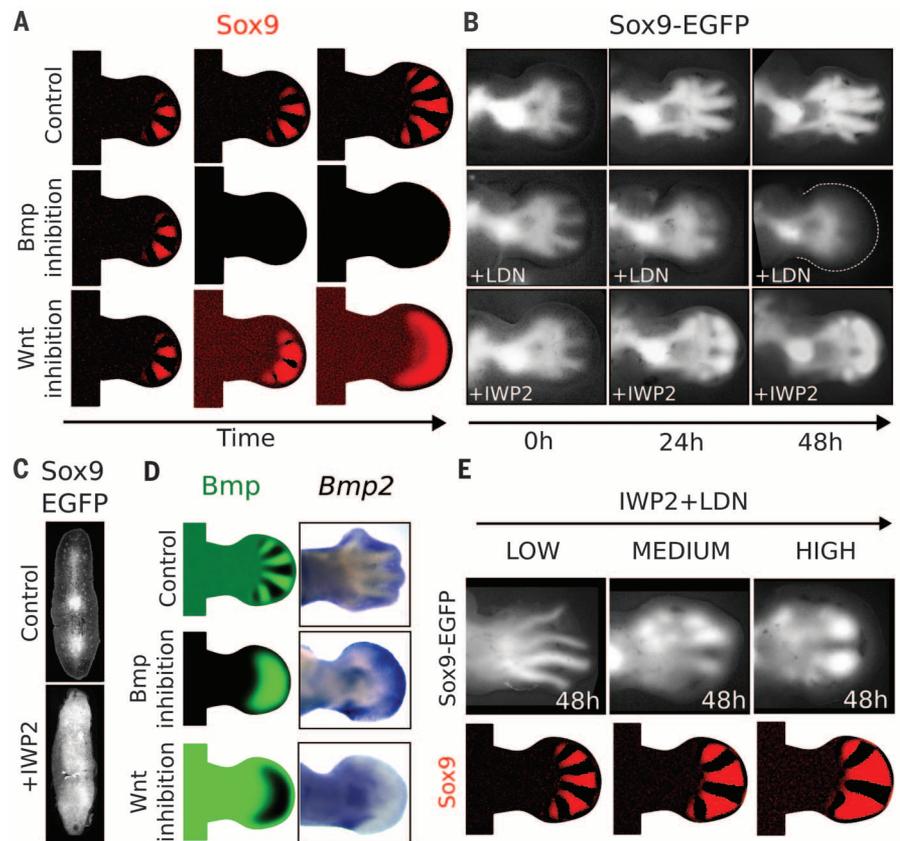
**Fig. 2. Sox9 is out of phase with BMP and WNT.** (A) Time course of *Bmp2* expression, Sox9-EGFP fluorescence (Sox9), BMP signaling (pSMAD), and WNT signaling activity (act- $\beta$ cat). Digit 4 (white arrowhead) appears first (E10.5), followed by digit 3 and then 2 (see Sox9). *Bmp2* and act- $\beta$ cat are out-of-phase with Sox9, whereas pSMAD is in-phase (see also fig. S2). (B) Top: Graphs show the progressive patterning along the AP axis measured from the images above. Black arrowheads point to digit 4. Bottom: A 1D simulation of the BSW model produces a Sox9 pattern that is out-of-phase with *Bmp* and *Wnt*, as shown in the experimental data above. (C) Network topology and equations of the BSW model.



**Fig. 3. Realistic computer simulation of digit patterning.** (A) When the BSW model was simulated inside an experimental limb growth map (left), Sox9 (red) formed a pattern with randomly oriented stripes biased along the PD axis (right). (B) Experimental *Hoxd13* expression was mapped into the growing model (left, heat-color map: blue = 0 and red = 1). With the *Hoxd13* modulation, the model creates a digit-like pattern (right), which eventually shows digit bifurcation (arrowhead). (C) *Fgf* expression in the AER was mapped into the model (solid line for *Fgf8*, dashed line for *Fgf4-9-17*) and used to simulate an Fgf signaling gradient (left, heat-color map: blue = 0 and red = 1). With the Fgf modulation, the model predicted a radially oriented Sox9 pattern with bigger wavelength toward the distal tip (right). (D) When Fgf and Hoxd13 jointly modulate the parameters  $k_4$  and  $k_7$ , the system shifts (dashed arrow) into the Turing space (gray region). (E) The simulated Sox9 pattern recapitulates the main features of (F) the experimental Sox9 expression in the digits, outlined by the white dotted lines. Arrowheads mark the late appearance of digits 5 and 1.



**Fig. 4. Model predictions and experimental perturbations.** (A) Sox9 patterns (red) predicted by the BSW model in different perturbations. Bmp inhibition causes loss of digital expression (second row), whereas Wnt inhibition predicts gradual loss of interdigits (third row). (B) Experimental manipulation of E11.5 limb buds in Sox9-EGFP limb culture confirmed the model predictions. Addition of the BMP signaling inhibitor LDN caused a loss of digital Sox9-EGFP (second row), whereas addition of the WNT inhibitor IWP2 caused the opposite result: Each digit expanded until complete fusion (third row). (C) Transverse sections show that WNT inhibition (+IWP2) induces expansion of Sox9-EGFP in the interdigital tissue (not only under the ectoderm). (D) The BSW model predicts that distal Bmp should be up-regulated and down-regulated in Bmp or Wnt signaling inhibition, respectively; the same behavior was observed for *Bmp2* expression in culture. (E) The model predicts progressively larger and fewer digits when both pathways are inhibited simultaneously (bottom row). Similar results are obtained by adding LDN and IWP2 simultaneously with progressively higher concentrations (low, medium, high) (upper row).



or out-of-phase with Sox9 (fig. S1). The analysis revealed in-phase and out-of-phase expression patterns for genes in the BMP and WNT pathways, but not for the FGF pathway. Specifically, *Bmp2* showed a pattern that is out-of phase with *Sox9* at all stages of digit patterning (Fig. 2A and fig. S1). None of the WNT ligands were differentially expressed, but target genes (e.g., *Axin2* and *Left1*) were expressed in Sox9<sup>-</sup> cells at E12 (fig. S1). This is consistent with the observation that WNT ligands produced in the ectoderm repress *Sox9* and chondrogenesis (20, 21). Indeed, knockout of  $\beta$ -catenin in the limb results in the expansion of *Sox9* toward the ectoderm, whereas  $\beta$ -catenin gain-of-function results in *Sox9* down-regulation (22, 23).

For a direct analysis of BMP, WNT, and FGF signaling activity, we performed immunohistochemistry for phosphorylated Smad1/5/8 (pSMAD), nonphosphorylated active  $\beta$ -catenin (act- $\beta$ cat), and phosphorylated extracellular signal-regulated kinase (pERK) on serial cryosections of Sox9-EGFP limbs. This allowed us to correlate the activity of the three pathways with the spatial dynamics of digit patterning. We found that BMP signaling (pSMAD) is high in the digits and correlates with Sox9-EGFP (Fig. 2A and fig. S2), supporting the positive effect of BMP signaling on *Sox9* (24–26) and its importance for digit patterning (27). In contrast to pSMAD, WNT signaling (act- $\beta$ cat) is high in the interdigital regions and has a pattern that is out-of-phase with Sox9 at all stages (Fig. 2A and fig. S2). This close spatial correspondence supports Wnt/ $\beta$ -catenin involvement in the Turing mechanism and is consistent with the repression of *Sox9* by Wnt/ $\beta$ -catenin (21, 22). Finally, FGF signaling (pERK) shows no periodic pattern but only the expected proximo-distal (PD) gradient, which is highest under the apical ectodermal ridge (AER) (fig. S2). Indeed, Fgf has previously been proposed as a modulator of the digit wavelength (9) rather than being part of the Turing network.

From the spatial relationships between these molecular patterns (Fig. 2B), we hypothesized that the Turing network that controls digit patterning is a three-node model implemented by Sox9, Bmp, and Wnt. Of the possible regulatory interactions between these three nodes, two are clear from the literature: the positive influence of BMP signaling on *Sox9* (24–26) and the inhibition of *Sox9* by WNT signaling (21–23). With these two interactions fixed, we performed a linear stability analysis for a three-component reaction-diffusion system (see supplementary text 2.2 and 2.3). Among all possible cases, we identified the topology shown in Fig. 2C (the Bmp-Sox9-Wnt, or BSW, model) as the simplest and most robust Turing network that can self-organize to produce a periodic pattern in which Sox9 is out-of-phase with both Bmp expression and Wnt signaling activity (Fig. 2B).

When simulated in 2D, the BSW model produces a pattern of randomly oriented stripes reminiscent of the pattern formed in micromass culture (fig. S3). The next challenge was to show that the model can reproduce the time course of

*Sox9* expression in a realistic simulation of limb development. Because Turing patterns are influenced by growth and shape changes of the underlying tissue (28, 29), we used a 2D growth model derived from clonal fate-mapping experiments (30). Within this tissue-movement map, the BSW network generated a stripy pattern that tended to orient along the PD axis (Fig. 3A, movie S4, and supplementary text 2.9.1). Next, because digit patterning appears to be relatively independent from the rest of the skeleton (9, 24, 27, 31), we chose to modulate the BSW model to generate the Turing instability only in the digital region defined by *Hoxd13* expression (Fig. 3B, movie S5, and supplementary text 2.9.2). This naturally promoted the sequential appearance of PD-oriented stripes and greatly reduced the variability of the pattern across different simulation runs (supplementary text 2.9.2); however, digit bifurcations occurred after E12 (Fig. 3B). We previously showed that distal *Hox* genes and FGF signaling jointly modulate the wavelength (digit width) to specify digit number and avoid bifurcations (9). When we combined modulation by FGF signaling (Fig. 3C and movie S6) and *Hoxd13* (Fig. 3D), the model robustly produced Sox9 patterns that recapitulated the *Sox9* expression in wild-type limbs (Fig. 3, E and F; fig. S4; supplementary text 2.9.4 and 2.12; and movies S7 and S8).

Finally, we used E11.5 Sox9-EGFP limb cultures (fig. S5) to test predictions from the computer model and to verify the Turing network. Reducing Bmp signaling in the simulation (from E11.5 onward) predicted a loss of digital Sox9 expression (Fig. 4A and movie S9) similar to the result of inhibiting BMP signaling in limb culture (Fig. 4B). Conversely, reducing Wnt signaling in the model predicted that all digits should expand until they merge into a continuous domain of Sox9 expression (Fig. 4A and movie S10). The same dynamics were observed when WNT signaling was inhibited in vitro (Fig. 4B). Our results reveal that upon abrogation of Wnt/ $\beta$ -catenin signaling, Sox9 expands not only in the sub-ectodermal mesenchyme (23) but also into the interdigital regions (Fig. 4C), proving that Wnt/ $\beta$ -catenin signaling is a necessary component of the periodic digital patterning process. The model also predicted that in all cases, Bmp expression should be opposite to that of Sox9, and we observed the same for *Bmp2* expression in the experiments (Fig. 4D). A further series of experiments, in which we implanted protein-soaked beads into the limb bud mesenchyme (fig. S6), also confirmed the predictions of the model and allowed us to hypothesize that the primary ligands are BMP2 and WNT3 (supplementary text 2.11.5 and movies S11 to S14).

Because inhibiting Bmp caused a loss of digits, whereas inhibiting Wnt caused a loss of interdigits, we explored whether the simultaneous inhibition of both pathways could balance out and instead cause an alteration of the pattern. The computer simulation predicted that a combination of both drugs would increase the digit period (the effective wavelength), resulting in an oligodactyly phenotype with fewer digits

(Fig. 4E and supplementary text 2.11.4). Application of both drugs to the limb bud culture confirmed this prediction, showing that the digit period becomes larger as the concentrations of both drugs increase (Fig. 4E and fig. S7). This supports both our BSW model and the general concept of a Turing-type process, because the new pattern was not just a down-regulation or up-regulation of *Sox9* but a genuine rearrangement of the already initiated periodic digital pattern.

Our study reveals the main feedbacks involved in digit patterning; however, other regulatory pathways may be involved to confer extra redundancy or robustness to the system. We propose that Wnt/ $\beta$ -catenin is essential for digit patterning and in particular for repressing *Sox9* in the interdigital regions. This suggests a possible mechanistic link between molecular patterning and mesenchymal condensations, as Wnt/ $\beta$ -catenin signaling may prevent the formation of  $\beta$ -catenin-cadherin complexes, which otherwise promote tighter cell adhesion (32). Finally, our model suggests a correlation between three separate functions of FGF signaling: control of growth (33), promotion of distal PD markers as *Hoxd13* (34), and promotion of larger digit wavelength (9). Taken together, this suggests that any FGF-driven change in limb bud size would naturally be accompanied by a change in PD digit length (*Hoxd13* domain) and in AP digit width (Turing wavelength) (fig. S8 and supplementary text 2.9.4), thereby implementing a perfect strategy for scale invariance.

## REFERENCES AND NOTES

1. D. Summerbell, J. H. Lewis, L. Wolpert, *Nature* **244**, 492–496 (1973).
2. R. D. Riddle, R. L. Johnson, E. Laufer, C. Tabin, *Cell* **75**, 1401–1416 (1993).
3. R. Zeller, J. López-Ríos, A. Zuniga, *Nat. Rev. Genet.* **10**, 845–858 (2009).
4. E. Zwilling, *Dev. Biol.* **9**, 20–37 (1964).
5. Y. Litingtung, R. D. Dahm, Y. Li, J. F. Fallon, C. Chiang, *Nature* **418**, 979–983 (2002).
6. A. M. Turing, *Bull. Math. Biol.* **52**, 153–197, 153–197 (1990).
7. S. A. Newman, H. L. Frisch, *Science* **205**, 662–668 (1979).
8. A. Gierer, H. Meinhardt, *Kybernetik* **12**, 30–39 (1972).
9. R. Shethi et al., *Science* **338**, 1476–1480 (2012).
10. A. Badugu, C. Kraemer, P. Germann, D. Menshykau, D. Iber, *Sci. Rep.* **2**, 991 (2012).
11. H. G. Hentschel, T. Glimm, J. A. Glazier, S. A. Newman, *Proc. Biol. Sci.* **271**, 1713–1722 (2004).
12. C. M. Leonard et al., *Dev. Biol.* **145**, 99–109 (1991).
13. T. Miura, K. Shiota, *Dev. Dyn.* **217**, 241–249 (2000).
14. R. Bhat et al., *BMC Dev. Biol.* **11**, 6 (2011).
15. V. Georgiadis et al., *Dev. Dyn.* **236**, 1014–1024 (2007).
16. E. Wright et al., *Nat. Genet.* **9**, 15–20 (1995).
17. Y. Nakamura et al., *Nat. Commun.* **2**, 251 (2011).
18. W. M. Kulyk, J. L. Franklin, L. M. Hoffman, *Exp. Cell Res.* **255**, 327–332 (2000).
19. H. Akiyama, M. C. Chaboissier, J. F. Martin, A. Schedl, B. de Crombrughe, *Genes Dev.* **16**, 2813–2828 (2002).
20. K. Summerhust, M. Stark, J. Sharpe, D. Davidson, P. Murphy, *GEP* **8**, 331–348 (2008).
21. D. ten Berge, S. A. Brugmann, J. A. Helms, R. Nusse, *Development* **135**, 3247–3257 (2008).
22. T. P. Hill, D. Später, M. M. Taketo, W. Birchmeier, C. Hartmann, *Dev. Cell* **8**, 727–738 (2005).
23. T. P. Hill, M. M. Taketo, W. Birchmeier, C. Hartmann, *Development* **133**, 1219–1229 (2006).
24. A. Bandyopadhyay et al., *PLoS Genet.* **2**, e216 (2006).
25. Q. Pan et al., *J. Cell. Physiol.* **217**, 228–241 (2008).
26. B. K. Zehentner, C. Dony, H. Burtscher, *J. Bone Miner. Res.* **14**, 1734–1741 (1999).

27. J. D. Bénazet *et al.*, *Development* **139**, 4250–4260 (2012).  
 28. E. J. Crampin, E. A. Gaffney, P. K. Maini, *Bull. Math. Biol.* **61**, 1093–1120 (1999).  
 29. T. Miura, K. Shiota, G. Morriss-Kay, P. K. Maini, *J. Theor. Biol.* **240**, 562–573 (2006).  
 30. L. Marcon, C. G. Arqués, M. S. Torres, J. Sharpe, *PLOS Comput. Biol.* **7**, e1001071 (2011).  
 31. C. Tickle, B. Alberts, L. Wolpert, J. Lee, *Nature* **296**, 564–566 (1982).  
 32. M. Bienz, *Curr. Biol.* **15**, R64–R67 (2005).  
 33. G. R. Martin, *Genes Dev.* **12**, 1571–1586 (1998).  
 34. N. Mercader *et al.*, *Development* **127**, 3961–3970 (2000).

## ACKNOWLEDGMENTS

The Sox9-EGFP mouse line was kindly provided by H. Akiyama and P. Jay. The mRNA probe for Bmp2 was kindly provided by C. Pujades. We thank R. Zeller for the Smad4 mutant limbs; A. R. Moreno for in situ of Fgf4, 8 and Twist1; and D. Barcena for the help with the micromass cultures. J.S. was supported by CRG and ICREA, L.R. by Ministerio de Economía y Competitividad (MINECO) BFU2010-16428, and Centro de Excelencia Severo Ochoa, SEV-2012-0208, and L.M. by the CRG; J.R. was funded by a Ph.D. Fellowship from the Fundação para a Ciência e Tecnologia. Confocal microscopy was done in the CRG Advanced Light Microscopy Unit, and FACS and microarrays in the CRG Core Facilities. The

microarray data discussed in this publication are accessible through Gene Expression Omnibus series accession number GSE518158.

## SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6196/566/suppl/DC1  
 Materials and Methods  
 Supplementary Text  
 Figs. S1 to S8  
 References (35–87)  
 Movies S1 to S14

10.1126/science.1252960

## HIV LATENCY

# Proliferation of cells with HIV integrated into cancer genes contributes to persistent infection

Thor A. Wagner,<sup>1,2\*</sup> Sherry McLaughlin,<sup>1,2\*</sup> Kavita Garg,<sup>3</sup> Charles Y. K. Cheung,<sup>3</sup> Brendan B. Larsen,<sup>2</sup> Sheila Styrchak,<sup>1</sup> Hannah C. Huang,<sup>1</sup> Paul T. Edlefsen,<sup>2,3</sup> James I. Mullins,<sup>2\*</sup> Lisa M. Frenkel<sup>1,2\*,†</sup>

Antiretroviral treatment (ART) of HIV infection suppresses viral replication. Yet if ART is stopped, virus reemerges because of the persistence of infected cells. We evaluated the contribution of infected-cell proliferation and sites of proviral integration to HIV persistence. A total of 534 HIV integration sites (IS) and 63 adjacent HIV *env* sequences were derived from three study participants over 11.3 to 12.7 years of ART. Each participant had identical viral sequences integrated at the same position in multiple cells, demonstrating infected-cell proliferation. Integrations were overrepresented in genes associated with cancer and favored in 12 genes across multiple participants. Over time on ART, a greater proportion of persisting proviruses were in proliferating cells. HIV integration into specific genes may promote proliferation of HIV-infected cells, slowing viral decay during ART.

Despite suppression of viral replication during ART, HIV reservoirs, measured by the number of resting CD4<sup>+</sup> T cells with infectious virus induced in cell culture, decline slowly (1). The mechanisms hypothesized to allow infectious proviruses to persist include long-lived latently infected cells (2); low-level HIV replication (2), potentially due to insufficient intracellular drug concentrations (3–5); and the proliferation of HIV-infected cells (2, 6–10). During ART, subpopulations of cells with identical HIV sequences comprise a progressively larger proportion of the persisting viral genomes (8), suggesting that proliferation of infected cells helps maintain the HIV reservoir. To further evaluate the contribution of infected-cell proliferation to HIV persistence, we developed a method [integration site loop amplification (ISLA)] to define sites of HIV integration in single cells and sequence up to 2.8 kb of the 3' region of the viral genome adjacent to the integration site, allowing us to link specific viral variants to specific integration sites (fig. S1).

A total of 534 proviral integration sites were sequenced from three participants (B1, L1, and R1) at three time points each: after 1 to 2.3, 4.1 to 8.2, and 11.3 to 12.7 years of suppressive ART (Fig. 1, A to C). HIV integration at the same chromosomal site was found in multiple cells within each participant throughout follow-up, whereas no identical integration sites were shared by different participants, suggesting that HIV-infected cells proliferate, as reported (2, 6–8).

The hypothesis was further investigated by comparisons of the viral genomes in 63 integration sites. When HIV C2V5 *env* sequences shared a specific integration site, the *env* sequences [~625 base pairs (bp)] were identical ( $n = 31$  C2V5 *env* sequences from 13 integration sites, except for one pair of sequences with a 1-bp difference). In contrast, among proviruses integrated at different positions in the human genome ( $n = 45$  unique integration sites), C2V5 *env* sequences were distinct except for three groups from participant B1 (fig. S2) (13 out of 13 versus 3 out of 45;  $P < 0.0001$ ). Sequences of the entire C2env-nef-3'LTR (long terminal repeat) region (~2.8 kb) from these three groups of proviruses had 2 to 39 nucleotide differences, indicating that many were distinct viruses (fig. S2). In comparison, eight pairs of C2env-nef-3'LTR

sequences with identical integration sites differed by 0 to 2 bases (mean 0.9), an error rate consistent with misincorporations during the ISLA protocol. Phylogenetic analyses of 396 additional HIV *env* sequences (8), together with those from the ISLA method, revealed multiple additional identical viral sequences (8), strongly suggesting that multiple HIV-infected clonal cell populations persist during ART (Fig. 2 and figs. S2 and S3).

Three approaches were used to explore whether the distribution of HIV integration sites observed was random or shaped by selective forces. Specifically, we examined the distribution of HIV integrations in genes associated with cancer, regulation of cell proliferation, or cell survival.

First, proviral integration sites were examined for overrepresentation in cancer-associated genes from five combined sources ( $n = 1332$  unique genes) (11–15). Across the three participants, 12.5% (36 out of 288) of the unique genes from all proviral integrations were annotated as associated with cancer (11–15), compared with only 5.19% (1332 out of 25,660) of the human genes in the human genome ( $P < 0.0001$ ). In addition, unique integrations in proliferating HIV-infected cells (defined as identical integration sites derived from  $\geq 2$  separate cells) (table S1, A to C) were also increased in cancer-associated genes (6 out of 34, 17.65% versus 1332 out of 25,660, 5.2%;  $P = 0.0076$ ), which suggests that HIV integration could disrupt the regulation of these genes, as is known to occur during tumor induction by nonacute oncogenic retroviruses (16).

Second, given that HIV integrates preferentially into actively transcribed genes, especially those activated upon HIV infection (17–19), we compared our participants' integration sites to the >44,000 integration sites mapped in acutely infected CD4<sup>+</sup> T cells (Jurkat cells). The majority of our participants' proviruses were integrated within genes (425 out of 534, 79.6%) (Fig. 1, A to C), particularly within introns (337 out of 425, 79.1%), as previously observed (9, 17, 20). The frequency of unique proviral integration sites in cancer-associated genes (12.70%) was similar to those mapped in the acutely infected Jurkat cells (11.14%) (21). However, our participants' integration sites in proliferating cells were enriched for cancer-associated genes over samples of the same size from the Jurkat cell data set ( $P = 0.0486$ ). In the Jurkat cell data set, 11.14% of integrations were found in cancer genes, compared with 15.97% in our participants ( $P = 0.0828$ ).

<sup>1</sup>Seattle Children's Research Institute, 1900 9th Avenue, Seattle, WA 98101, USA. <sup>2</sup>University of Washington, Seattle, WA, USA.

<sup>3</sup>Fred Hutchinson Cancer Research Center, Seattle, WA, USA.

\*These authors contributed equally to this work. †Corresponding author. E-mail: lfrenkel@uw.edu



## Digit patterning is controlled by a Bmp-Sox9-Wnt Turing network modulated by morphogen gradients

J. Raspopovic *et al.*  
*Science* **345**, 566 (2014);  
DOI: 10.1126/science.1252960

*This copy is for your personal, non-commercial use only.*

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

**The following resources related to this article are available online at [www.sciencemag.org](http://www.sciencemag.org) (this information is current as of April 10, 2016 ):**

**Updated information and services**, including high-resolution figures, can be found in the online version of this article at:  
</content/345/6196/566.full.html>

**Supporting Online Material** can be found at:  
</content/suppl/2014/07/30/345.6196.566.DC1.html>

A list of selected additional articles on the Science Web sites **related to this article** can be found at:  
</content/345/6196/566.full.html#related>

This article **cites 80 articles**, 28 of which can be accessed free:  
</content/345/6196/566.full.html#ref-list-1>

This article has been **cited by** 13 articles hosted by HighWire Press; see:  
</content/345/6196/566.full.html#related-urls>

This article appears in the following **subject collections**:  
Development  
</cgi/collection/development>