doi:10.1038/nature08838

LETTERS

A computational model of teeth and the developmental origins of morphological variation

Isaac Salazar-Ciudad[1,2] & Jukka Jernvall[2,3]

The relationship between the genotype and the phenotype, or the genotype–phenotype map, is generally approached with the tools of multivariate quantitative genetics and morphometrics[1–4]. Whereas studies of development[5–7] and mathematical models of development[4,8–12] may offer new insights into the genotype–phenotype map, the challenge is to make them useful at the level of microevolution. Here we report a computational model of mammalian tooth development that combines parameters of genetic and cellular interactions to produce a three-dimensional tooth from a simple tooth primordia. We systematically tinkered with each of the model parameters to generate phenotypic variation and used geometric morphometric analyses to identify, or developmentally ordinate, parameters best explaining population-level variation of real teeth. To model the full range of developmentally possible morphologies, we used a population sample of ringed seals (*Phoca hispida ladogensis*)[13]. Seal dentitions show a high degree of variation, typically linked to the lack of exact occlusion[13–16]. Our model suggests that despite the complexity of development and teeth, there may be a simple basis for dental variation. Changes in single parameters regulating signalling during cusp development may explain shape variation among individuals, whereas a parameter regulating epithelial growth may explain serial, tooth-to-tooth variation along the jaw. Our study provides a step towards integrating the genotype, development and the phenotype.

The relatively well-studied mammalian tooth is suitable for identifying the developmental principles linking genotype to phenotype. Cusps constitute the major morphological features of teeth, and cusp shape, size, position and number are used to differentiate taxa. At the onset of each cusp formation—non-dividing epithelial signalling centres—the secondary enamel knots appear[17]. Experimental evidence from mouse molars implicates molecular signalling in activating and inhibiting the formation of enamel knots[11,18], and computational modelling of these dynamics can account for basic aspects of tooth formation[19]. Previous models[19,20], however, approximate tooth shapes without detailed implementation of the mechanical properties of cells. Because mechanical forces have a central role in tissue architecture[20–23], here we have constructed a model integrating experimentally inferred genetic interactions with tissue dynamics (Fig. 1a), thereby allowing us to study population-level multivariate variation.

Signalling parameters of the model numerically represent how strongly specific growth factors affect expression of one another and their diffusion (see Methods). The gene network topology is on the basis of empirical evidence from mouse molar development (Fig. 1a and Methods). This includes BMPs as differentiation inducers[18] and FGF10 in the regulation of epithelial proliferation[24]. Signalling is integrated with tissue dynamics by how strongly specific growth factors affect proliferation. Cell shape is affected by mechanical forces caused by tissue growth, and by parameters such as adhesion and repulsion (Fig. 1a, b and Methods). The formation of each enamel knot leads to lateral inhibition preventing the formation of new enamel knots in the immediate vicinity, which in turn leads to the folding of the epithelium owing to spatial differences in cell proliferation (Fig. 1b). Virtual, *in silico*, tooth shape is a product of a dynamic process in which each parameter combination determines only the strengths of interactions between gene products and cells (Fig. 1). Therefore, and conforming to the current evidence on tooth development[17,20], a specific tooth shape does not require cusp specific genes or a coordinate system.

We used the model to examine variation in the teeth of phocid seals (Fig. 2a). Seal dentitions provide a modest test of the generality of the model and have traits that make them suitable for studying population-level variation. The postcanine dentition of phocid seals is relatively simple; effectively two-dimensional due to the lack of

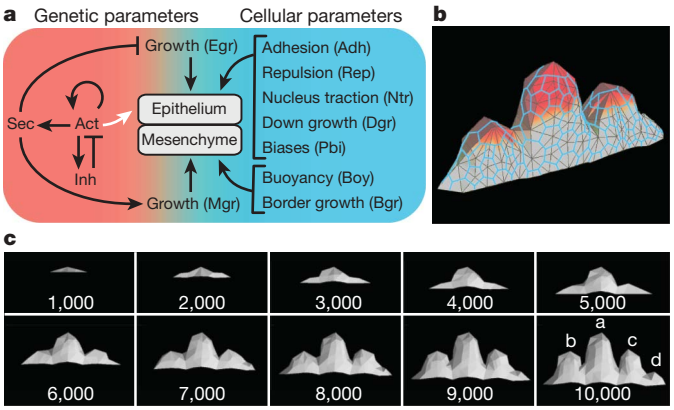

Figure 1 | A model integrating gene networks and tissue mechanics. a, Nine parameters regulating gene network properties and ten parameters regulating cellular properties integrate the behaviour of cells. Act induces the differentiation of epithelial cells into enamel knots (white arrow), which results in the production of Inh and also Sec regulating tissue growth. **b**, Tissue morphology is modelled for the cells of the inner enamel epithelium. The underlying mesenchyme is a three-dimensional space in which molecules and mechanical stresses diffuse. Molecular diffusion between the immediate surroundings of two epithelial cells is calculated using Fick's law of diffusion (diffusing inhibitor shown in colour). A triangular mesh connects cells centres (black lines) and the position of each corner is calculated as a Voronoi node (blue mesh). Cell shape depends on the positions and number of neighbouring cells. **c**, The initial conditions consists of seven epithelial cells representing the tip of the oral epithelium invagination and developing tooth shape can be visualized at any time point, here shown at 1,000 time intervals. At each interval, all equations are integrated using the Euler method and all variation simulations used 10,000 time points. Anterior towards the left. a–d denote cusp names used to identify seal tooth cusps.

[1]Departament de Genètica i Microbiologia, Facultat de Biociències, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain. [2]Developmental Biology Program, Institute of Biotechnology, University of Helsinki, PO Box 56, FIN-00014 Helsinki, Finland. [3]Department of Ecology and Evolution, Stony Brook University, Stony Brook, New York 11794, USA.

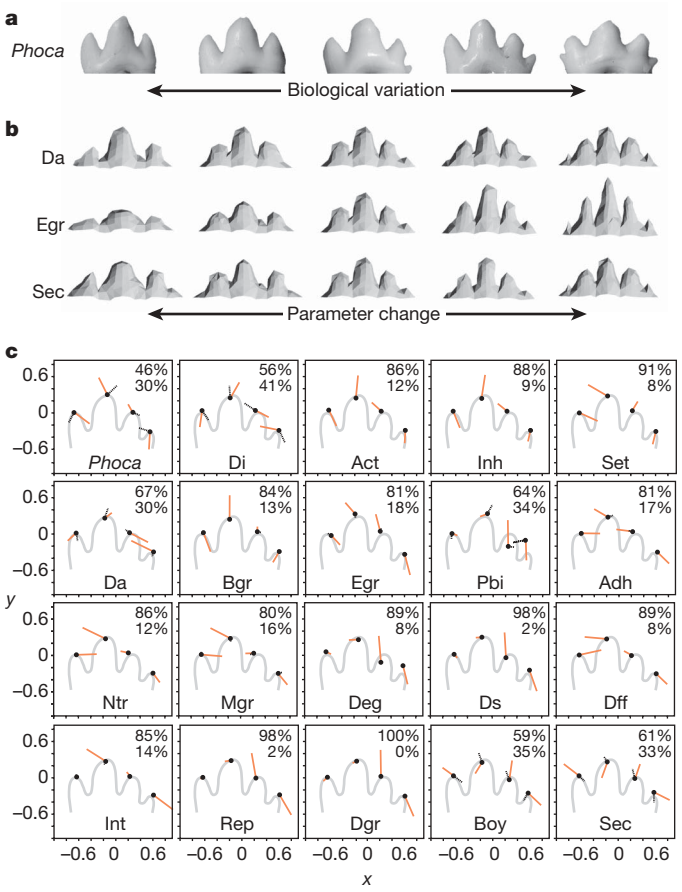

Figure 2 | Shape of variation in real and *in silico* seal teeth. **a**, The fourth postcanine tooth of Lake Ladoga seals varies from three to five cusped shapes. **b**, Systematical tinkering with each model parameter (shown at 20% intervals for the three examples) produces variation in cusp shape, size and number. **c**, Component loadings based on procrustes superimposition (variance-covariance matrix) and principal component analyses of the four tallest cusp positions (teeth with four and five cusps, $n = 67$) and of each simulated tooth population after changing each parameter ($n = 22$ to 99 depending on how many teeth had at least four cusps). Loadings shown for the first two components together with percentages of variance explained. The whiskers show the direction and relative strength that the first two components affect variation (the whisker lengths of the second components, dashed black line, are scaled based on variance explained relative to the first component, orange line). The grey outline shows the mean cusp pattern of the real teeth (*Phoca*). Anterior towards the left. For parameter names, see Methods Summary.

lateral cusp position variation[13], and lacks the refined occlusion typical of most extant mammals. These relaxed functional demands are reflected in the high degree of dental variation found in seal populations[13–16], and seal teeth should provide a relatively comprehensive sample of developmentally possible phenotypes. First we studied a sample ($n = 70$) of Lake Ladoga seals (*Phoca hispida ladogensis*), a postglacial landlocked population of ringed seals. The postcanine teeth of ringed seals consists typically of four cusps aligned in a row, but three and five cusped teeth are also frequently present (Fig. 2a)[13]. The seal dentition is derived from a more complex carnivoran pattern[25] and the tooth morphology of ringed seals is reminiscent of the pretribosphenic triconodont pattern in mammalian evolution[26].

Our initial exploration of the model showed that it produces virtual teeth resembling real seal teeth (Fig. 1). Moreover, the developmental sequence of cusps of the virtual teeth (Fig. 1c) corresponds to the order of cusp development in real seals[27]. Next we systematically tinkered with the parameters to produce a virtual tooth that closely resembles an average real fourth lower postcanine (P4), in terms of cusp number (four) and the angle of three tallest cusps (top-cusp angle, ref. 13). This virtual 'wild type' seal tooth was then mutated by individually changing the nine genetic and ten cellular parameters (Fig. 2b and Supplementary Fig. 2). We changed each parameter at 2% intervals in both directions from the wild-type value (up to ±98%). Variation in tooth shapes in the 19 virtual and one real seal tooth populations was then analysed individually using geometric morphometrics. Because cusp tips of real and virtual teeth show little variation in the third dimension[13], we report the results analysing cusp tip positions in two dimensions from a lateral view. We measured the x–y positions of the four main cusps (cusps a, b, c and d) in all four- and five-cusped seal teeth ($n = 22$ to 99) and carried out procrustes superimposition of each tooth population followed by principal component analysis (ref. 13, Methods and Supplementary Information).

The results show that the first two principal components of real teeth account for 46% and 30% of total variance, respectively (Fig. 2c). The first component mainly affects how blunt the cusp configuration is, whereas the second component also influences the relative position of cusps anterio-posteriorly (Fig. 2c). To compare the overall differences in the shape of variation, or developmentally ordinate phenotypic variation, we tabulated the principal component loadings of the first two components for each cusp, and plotted the difference between real teeth and virtual teeth (Fig. 3a). The results show (Fig. 3a and Supplementary Table 4) that the first five parameters producing the most realistic variation are activator self-regulation (Act), activator diffusion (Da), inhibitor strength (Inh), inhibition diffusion (Di), and secondary signal threshold (Set). Of these, the activator and inhibitor parameters are central in the formation of enamel knots and we interpret them to constitute a patterning kernel of tooth cusps. We note that activation diffusion and inhibition diffusion affect both the first two principal components (Fig. 2c), suggesting that these parameters can individually account for a substantial portion of the variation found in real teeth.

Next we examined how different parameters produce variation in cusp number in addition to cusp position. The top-cusp angle and cusp number are linked in seal teeth; larger angles, hence blunter cusp

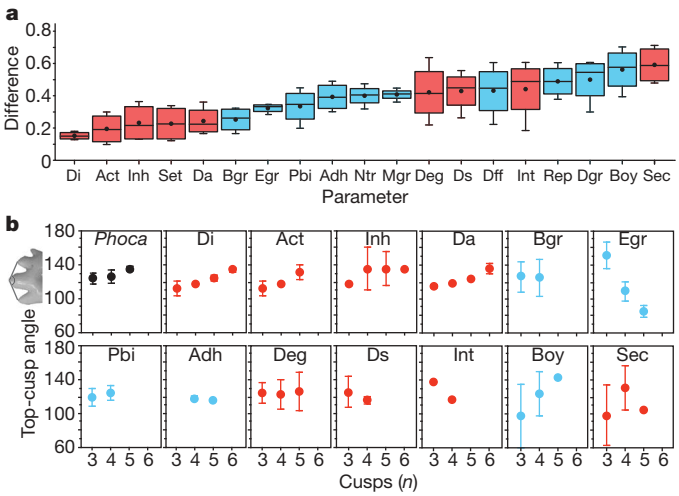

Figure 3 | Variation in cusp position and number implicate the same patterning kernel parameters. **a**, Differences in the first two principal component loadings (plotted for all four cusps) between real seal teeth and *in silico* teeth show that genetic parameters belonging to the activator–inhibitor loop regulating enamel knot formation produce the most realistic variation. Boxes enclose 50% of cusps; the median and the mean are indicated with a horizontal bar and a circle, respectively, and whiskers denote range. **b**, Tinkering systematically with each model parameter (shown at 20% intervals) produces variation in cusp number in 13 of the 19 parameters. We note the tendency for increasing top-cusp angle with increasing number of cusps in real seal teeth and *in silico* teeth varied with the patterning kernel parameters. Genetic parameters are in red, cellular parameters are in blue. Error bars denote s.d.

configurations, are associated with a larger number of cusps[13]. Developmentally, closer placement of enamel knots is thought to produce both more equal sized and a larger number of cusps[13]. The results show that only the patterning kernel parameters produce realistic changes both in cusp number and top-cusp angle (Fig. 3b), suggesting that these parameters are the best candidates for accounting for variation among individuals. An obvious assumption underlying these interpretations is that our wild-type *in silico* seal tooth has representative variational properties. To test how dependent the variational properties are on any given *in silico* seal tooth, we searched the parameter space and located nine other parameter combinations that produced seal tooth morphologies (see Supplementary Figs 3, 4 and Supplementary Table 5). For each of the nine wild-type seal teeth, we systematically tinkered with each of the model parameters and measured the top-cusp angles in relation to cusp number. The variation produced by the patterning kernel parameters remains more similar to the real teeth than variation produced by the other parameters (mean top-cusp angle difference to real teeth is 9.2 degrees for the patterning kernel teeth and 17.2 degrees for the remaining teeth, $P = 0.0007$; Mann–Whitney *U*-test, see Supplementary Information). In connection to changes in cusp number, we noticed that owing to changing allocation of cells to enamel knots, cusp c becomes narrower as the fourth cusp, cusp d, appears (Supplementary Fig. 5). This inverse relationship between cusps was also detected in real teeth (Supplementary Fig. 5), and suggests that in cladistic analyses, for example, relative cusp size and the presence of adjacent small conules may be non-independent characters.

In addition to individual variation, mammalian teeth show serial variation along the tooth row. In ringed seals this variation is a relatively subtle increase in the top-cusp angle towards the anterior teeth, whereas in grey seals (*Halichoerus grypus*) there is a heterodont switch from three-cusped to one-cusped teeth (Fig. 4a). Searches within the virtual teeth variation (in 2% intervals as before; see Supplementary Table 1) show that only the cellular parameter regulating epithelial growth (Egr) produces the ringed seal tooth row (Fig. 4b). We propose that our results on the different parameters underlying changes in individual tooth shape (Fig. 3) and changes along the tooth row (Fig. 4) reflect the decoupling of shape from serial variation. This decoupling, which can be postulated to have enabled the evolution of heterodonty, agrees with the notion of restricted pleiotropy being a solution to evolving complex traits[3]. Mechanistically, genetic parameters regulate lateral inhibition, which in turn affects cusp spacing, whereas Egr regulates the relative growth of epithelium over mesenchyme, which in turn affects pointedness of cusps (and only indirectly cusp spacing; Fig. 2b and Supplementary Fig. 6). Cusp pointedness is a central feature of mammalian heterodonty and perhaps the reason why the parameter decoupling has the observed polarity. In the grey seal, however, regulation of Egr may not be sufficient to produce its tooth row—the patterning kernel parameter Act is also required (Fig. 4b). Nevertheless, in both seal species, a constant tooth-to-tooth parameter change can produce tooth rows, even when morphological change is discontinuous as in grey seals (Fig. 4). We note that the constant change of parameter values along the tooth row may provide a mechanistic basis for aspects of the inhibitory cascade model determining mammalian molar proportions[11].

In conclusion, despite the complexity of development and teeth, a large proportion of variation can be explained by changes in single model parameters. Even though these results do not exclude the possibility of a complex, polygenic nature of tooth morphology, they are suggestive of a simple basis for dental variation. Because genes associated with tooth development seem to be largely conserved across vertebrates[28], it remains to be tested whether a similar patterning kernel, or 'genetic architectural stereotype'[29], underlies the regulation of multicuspid teeth beyond mammals. Considering the quantitative nature of traits, the inclusion of an allelic component into the model will complete the link between genotype and phenotype. In these kinds of analyses the inclusion of developmental models is valuable to detect the non-normal and discontinuous changes in the phenotype despite gradual changes in developmental parameters. These advances should eventually lead to the integration of the process of development with the genotype and the phenotype, something that could readily be recognized as the 'laws of correlation of growth'[30].

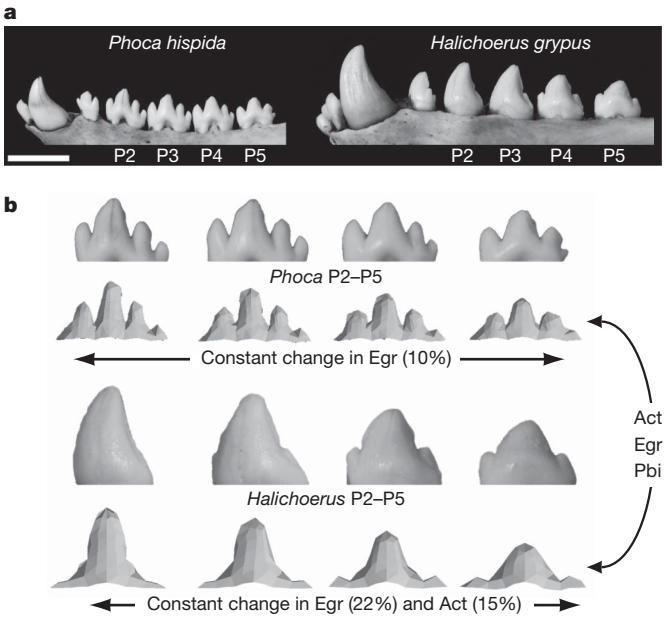

Figure 4 | Serial tooth-to-tooth variation implicates a cellular parameter. **a**, The four ringed seal (*Phoca*) postcanines (P2–P5) show a gradual change in morphology, whereas the grey seal (*Halichoerus*) tooth row has a transition from a three- to a one-cusped shape. **b**, Constant change in parameter Egr, the epithelial growth rate, produces change in relative cusp height similar to the morphological change observed in the ringed seal tooth row. We note the curvature of the *in silico* P2 main cusp, resembling that of real teeth. Three parameters (Act, Egr and Pbi) can be used to turn an *in silico* ringed seal tooth into a grey seal tooth, and constant changes in two parameters (Egr and Act) produce changes reminiscent of changes along the grey seal tooth row, despite the heterodont change in morphology. Grey seal was produced by manually tinkering with ringed seal parameters. Scale bar, 10 mm.

METHODS SUMMARY

The model (available with the source code at http://www.biocenter.helsinki.fi/bi/evodevo/isc/programpage.html) includes an initial grid of seven epithelial cells arranged in a hexagon and two layers of underlying mesenchymal cells (Methods). The epithelial grid grows by cell division and folds owing to forces arising from cells. Cell division occurs after cell size reaches a threshold. The same set of coupled differential equations governs gene product interactions in each cell. Some gene products diffuse between cells according to Fick's law (finite volume method). Act autoregulates itself and, after a threshold, causes the differentiation of an enamel knot and induces the production of Inh and secondary signals (Sec). Mechanical interactions of cells are regulated by the following genetically encoded parameters: cell elasticity to compression by other cells (cell repulsion, Rep), cells resistance to separation by a force proportional to the parameter (adhesion, Adh), attachment force between the nucleus and the cell borders (nucleus traction, Ntr), a default rate of epithelial growth along the triangular mesh (Egr), gradual cessation of cell signalling and growth (differentiation, Dff), differential growth of the anterior and posterior borders of teeth (growth bias, Pbi), tendency of the epithelial cells at the tooth borders (cervical loops) to engulf the underlying dental mesenchyme (cervical loop downward growth, Dgr), and mechanical resistance of the mesenchyme to the invagination of the epithelium (mesenchymal buoyancy, Boy). The landmarks for geometric morphometrics were obtained from digital images of real and virtual teeth using ImageJ (http://rsb.info.nih.gov/ij/) and all analyses were done using PAST (http://folk.uio.no/ohammer/past/index.html).

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

Received 21 July 2009; accepted 15 January 2010.
Published online 10 March 2010.

1. Roff, D. A. A centennial celebration for quantitative genetics. *Evolution* **61,** 1017–1032 (2007).
2. Zeng, Z. B. QTL mapping and the genetic basis of adaptation: recent developments. *Genetica* **123,** 25–37 (2005).
3. Wagner, G. P. *et al.* Pleiotropic scaling of gene effects and the 'cost of complexity'. *Nature* **452,** 470–472 (2008).
4. Polly, P. D. Developmental dynamics and G-matrices: can morphometric spaces be used to model phenotypic evolution? *Evol. Biol.* **35,** 83–96 (2008).
5. Prud'homme, B., Gompel, N. & Carroll, S. B. Emerging principles of regulatory evolution. *Proc. Natl Acad. Sci. USA* **104,** 8605–8612 (2007).
6. McGregor, A. P. *et al.* Morphological evolution through multiple *cis*-regulatory mutations at a single gene. *Nature* **448,** 587–590 (2007).
7. Miller, C. T. *et al. cis*-Regulatory changes in Kit ligand expression and parallel evolution of pigmentation in sticklebacks and humans. *Cell* **131,** 1179–1189 (2007).
8. Salazar-Ciudad, I., Newman, S. A. & Solé, R. V. Phenotypic and dynamical transitions in model genetic networks. I. Emergence of patterns and genotype-phenotype relationships. *Evol. Dev.* **3,** 84–94 (2001).
9. Harris, M. P., Williamson, S., Fallon, J. F., Meinhardt, H. & Prum, R. O. Molecular evidence for an activator-inhibitor mechanism in development of embryonic feather branching. *Proc. Natl Acad. Sci. USA* **102,** 11734–11739 (2005).
10. Evans, T. M. & Marcus, J. M. A simulation study of the genetic regulatory hierarchy for butterfly eyespot focus determination. *Evol. Dev.* **8,** 273–283 (2006).
11. Kavanagh, K. D., Evans, A. R. & Jernvall, J. Predicting evolutionary patterns of mammalian teeth from development. *Nature* **449,** 427–432 (2007).
12. Nahmad, M., Glass, L. & Abouheif, E. The dynamics of developmental system drift in the gene network underlying wing polyphenism in ants: a mathematical model. *Evol. Dev.* **10,** 360–374 (2008).
13. Jernvall, J. Linking development with generation of novelty in mammalian teeth. *Proc. Natl Acad. Sci. USA* **97,** 2641–2645 (2000).
14. Lönnberg, E. Några egendomliga variationer i tanduppsättningen hos vikaren, *Phoca hispida*. *Fauna och Flora* **18,** 116–126 (1923).
15. Cruwys, L. & Friday, A. Visible supernumerary teeth in pinnipeds. *Polar Rec.* **42,** 83–85 (2006).
16. Miller, E. H. *et al.* Variation and integration of the simple mandibular postcanine dentition in two species of phocid seal. *J. Mamm.* **88,** 1325–1334 (2007).
17. Jernvall, J., Keränen, S. V. E. & Thesleff, I. Evolutionary modification of development in mammalian teeth: Quantifying gene expression patterns and topography. *Proc. Natl Acad. Sci. USA* **97,** 14444–14448 (2000).
18. Kassai, Y. *et al.* Regulation of mammalian tooth cusp patterning by ectodin. *Science* **309,** 2067–2070 (2005).
19. Salazar-Ciudad, I. & Jernvall, J. A gene network model accounting for development and evolution of mammalian teeth. *Proc. Natl Acad. Sci. USA* **99,** 8116–8120 (2002).
20. Osborn, J. W. A model of growth restraints to explain the development and evolution of tooth shapes in mammals. *J. Theor. Biol.* **255,** 338–343 (2008).
21. Forgacs, G. & Newman, S. A. *Biological Physics of the Developing Embryo* (Cambridge Univ. Press, 2005).
22. Keller, R. Mechanisms of elongation in embryogenesis. *Development* **133,** 2291–2302 (2006).
23. Lecuit, T. & Lenne, P. F. Cell surface mechanics and the control of cell shape, tissue patterns and morphogenesis. *Nature Rev. Mol. Cell Biol.* **8,** 633–644 (2007).
24. Kettunen, P. *et al.* Associations of FGF-3 and FGF-10 with signaling networks regulating tooth morphogenesis. *Dev. Dyn.* **219,** 322–332 (2000).
25. Rybczynski, N., Dawson, M. R. & Tedford, R. H. A semi-aquatic Arctic mammalian carnivore from the Miocene epoch and origin of Pinnipedia. *Nature* **458,** 1021–1024 (2009).
26. Luo, Z.-X. Transformation and diversification in early mammal evolution. *Nature* **450,** 1011–1019 (2007).
27. Stewart, B. E., Innes, S. & Stewart, R. E. A. Mandibular dental ontogeny of ringed seals (*Phoca hispida*). *Mar. Mamm. Sci.* **14,** 221–231 (1998).
28. Fraser, G. J. *et al.* An ancient gene network is co-opted for teeth on old and new jaws. *PLoS Biol.* **7,** e1000031 (2009).
29. Buchanan, A. V., Sholtis, S., Richtsmeier, J. & Weiss, K. M. What are genes "for" or where are traits "from"? What is the question? *Bioessays* **31,** 198–208 (2009).
30. Darwin, C. *On the Origin of Species by Means of Natural Selection* 1st edn (Murray, 1859).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank I. Corfe, A. R. Evans, J. Fierst, M. Fortelius, B. Julia, S. Sova, J. Hakanen, E. Harjunmaa, N. Navarro, I. Thesleff, P. C. Wright and S. Zohdy for comments, discussions, and support on this work, A. Kangas for help in data collection, and M. Hildén and I. Hanski (Finnish Museum of Natural History) for access to collections. This study was supported by the Ramón y Cajal Program (RYC-2007-00149) and the Academy of Finland.

Author Contributions I.S.-C. and J.J. conceived the study; I.S.-C. constructed the computational model and performed computer simulations; J.J. obtained the empirical data; I.S.-C. and J.J. performed quantitative analyses and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to I.S.-C. (isaac.salazar@uab.cat) or J.J. (jernvall@fastmail.fm).

METHODS

The model incorporates gene network dynamics (for example, see refs 31, 32) and mechanical interactions of cells (for example, see refs 33–35), and therefore is not a statistical or geometric[36] description of variation. Morphology results from the dynamics of model and variation from differences in the parameters describing how strongly genes and cells interact. Specifically, the model covers tooth morphogenesis from the bud stage to the onset of mineralization (termination of cuspal morphogenesis). Model output is morphology, specified as the position of cells in three-dimensional space, and patterns of gene expression in three-dimensional space. The tooth shape is the epithelial–mesenchymal interface, and because seal teeth have thin enamel, this is a close proximation of the mineralized seal tooth.

Tissue morphology is modelled by using the cells of the inner enamel epithelium. Each epithelial cell has a centre and boundaries. Cell shape depends on the relative positions and number of neighbouring epithelial cells, typically forming irregular honeycomb tessellation. The boundary between two epithelial cells is quadrilateral with one of its sides towards the apical surface (facing the stelate reticulum) and one towards the basal surface (facing the mesenchyme). A triangular mesh connects cell centres and the position of each apical corner is calculated as a Voronoi node; each corner is equidistant from the three closest cell centres. Therefore, a cell encloses all the points in space that are closer to its centre than to the centre of any other cell. The basal cell corners are the same as the apical ones but one *z*-unit down. Cells can change neighbours through the addition of new cells by cell division, but not by cell migration. The underlying mesenchyme is a three-dimensional space in which molecules and mechanical stresses diffuse and spread.

Diffusion of molecules is possible between cells in the epithelium, between mesenchymal cells and between the mesenchyme and the epithelium but not into the stellate reticulum above the inner enamel epithelium. The dental epithelium acts, thus, as a wall of a diffusion chamber. Molecules leave the system from the epithelial borders and the mesenchyme. Molecular diffusion between the immediate surroundings of two cells is calculated using Fick's law of diffusion by making the molecular flux between two cells proportional to the area in which they are in contact or near each other (finite volume method). This method allows for accurate calculations even when cells change their shapes.

The gene network is modelled to capture the core molecular events leading to the formation of enamel knots. Act autoregulates itself and, past a threshold, causes the differentiation of an enamel knot and induces the production of Inh and Sec. In reality these model molecules may be encoded by several genes, and at several layers of regulatory machinery. Activators include molecular candidates such as *Bmp4*, *activin bA* (also known as *Inhba*) and many genes from the *Wnt* family, and inhibitors include at least *Shh* and *Sostdc1*. The secondary signals are growth factors that are upregulated in the enamel knots later and the main function of which seems to be the regulation of cell proliferation and differentiation rather than direct inhibition of new enamel knots. These include at least *Fgf4* and *Bmp2*. All equations, explained later, are integrated by the Euler numerical method. Here all model parameters are depicted as *k* plus a subscript with the three letters used to name the parameters in the main text.

Cell repulsion and adhesion. To model cells as physical entities, repulsion was implemented to prevent cells from getting too close to each other. A cell responds by exerting a force if the centre of a neighbouring cell is closer than the original distance separating them (for example, when new cells form by cell division). Between cell *i* and cell *j* we have:

$$\mathbf{f}_{ij} = k_{\mathrm{Rep}}\left(||\mathbf{p}_j - \mathbf{p}_i|| - ||\mathbf{p}_o||\right)(\mathbf{p}_j - \mathbf{p}_i) \tag{1}$$

Where k_{Rep} is the repulsion model parameter, a stiffness constant that describes how much cells respond to applied stresses. $\mathbf{p}_i$ is the position in three dimension of cell *i*. || denotes modulus and $||\mathbf{p}_o||$ is the original distance between two cells. Note that $||\mathbf{p}_o||$ is not a parameter of the model but results from the model dynamics (except in the initial conditions where it equals one). When the distance between two cell centres is larger than $||\mathbf{p}_o||$ there is a constant traction force resulting from the adhesion of cells. For these distances we have:

$$\mathbf{f}_{ij} = k_{\mathrm{Adh}}\frac{(\mathbf{p}_j - \mathbf{p}_i)}{\left\|\mathbf{p}_j - \mathbf{p}_i\right\|} \tag{2}$$

Where k_{Adh} is a model parameter specifying the adhesion force. Taken together, for each iteration, a cell position changes as follows:

$$\frac{\partial \mathbf{p}_i}{\partial t} = \sum_{\text{for all neighbours } j \text{ of } i} f_{ij} \tag{3}$$

Nucleus traction. After biomechanical studies[23], we also include a force attaching each cell centre to cell borders. This force tends to cause rounding of epithelial cells:

$$\frac{\partial \mathbf{p}_i}{\partial t} = k_{\mathrm{Ntr}}(1-d_i)\left(-\mathbf{p}_i + \frac{1}{n_i}\sum_{\text{for all neighbours } j \text{ of } i}\mathbf{p}_j\right) \tag{4}$$

Where $\mathbf{p}_j$ is the position of each neighbouring cell, n_i is the number of neighbours of cell *i* and k_{Ntr} is a model parameter describing how strong this force is. d_i is the state of differentiation of cell *i* (see later).

Epithelial growth and differentiation. Epithelial cells have a default rate of growth and cells grow by pushing the cell centres of their neighbours away along the triangular mesh. Thus, the displacement of a cell centre is in the direction of a unit vector arising by summing the unit vectors joining the cell centre with that of its neighbours and dividing the vector by its modulus (length). This vector is then multiplied by the model parameter k_{Egr}. Dff decreases growth and this is implemented by multiplying growth by $1 - d_i$:

$$\frac{\partial \mathbf{p}_i}{\partial t} = k_{\mathrm{Egr}}(1-d_i)\frac{\sum_{\text{for all neighbours } j \text{ of } i}\hat{\mathbf{u}}_{ij}}{\left\|\sum_{\text{for all neighbours } j \text{ of } i}\hat{\mathbf{u}}_{ij}\right\|} \tag{5}$$

Where $\hat{\mathbf{u}}_{ij}$ is the unit vector between a cell *i* and its neighbour *j*. The total displacement in each cell is $k_{\mathrm{Egr}}(1 - d_i)$. Note that the differentiation of cell *i*, d_i, ranges between 0 and 1. This is affected by a model parameter k_{Dff} which describes how strongly Sec promotes differentiation:

$$\frac{\partial d_i}{\partial t} = k_{\mathrm{Dff}}[\mathrm{Sec}] \tag{6}$$

Cells divide after the connection between two cell centres is equal to or exceed two units of space (initial distance between cells equals one unit) and is implemented by placing a new cell at the midpoint between the two mother cells. This gives qualitatively similar results to a more realistic, but computationally slower, implementation of cell division in which a cell is split into two. A new cell adds a new cell centre and connections to the original triangular mesh (Supplementary Fig. 1). The concentration of molecules and the differentiation state of the new cell is the average of the two mother cells. If one of the mother cells is an enamel knot cell, the new cell is not a knot cell. Any growth vector that is towards the stellate reticulum, which has an increased hydrostatic pressure, is inverted in respect to the plane of the epithelium. In addition, once formed, enamel knots are not allowed to move down (following empirical evidence[37,38]).

Growth biases. The anterior and posterior borders of developing teeth grow at different rates[17] and this was modelled by multiplying the displacement along the anterio-posterior axis (along the *y*-coordinates) in the anterior, by parameter k_{Abi}, and posterior, by parameter k_{Pbi}, borders of the tooth. The anterior border is defined as the border cells that in the initial condition are anterior to the cell in the centre (positive *y* values), and all their descendants that also lay in the teeth border. The posterior border is defined in the same way for negative values of *y*. Because of the symmetric nature of these biases, we only modulated k_{Pbi} in the variation analysis. To delay the cervical loop formation in anterior–posterior directions (as seen in real teeth), mesenchymal buoyancy parameter (k_{Bgr}) is used to multiply the *z*-value of border cells.

Cervical loop downward growth and mesenchymal growth. Epithelial cells at the border of the tooth (cervical loop) grow to engulf the underlying dental mesenchyme. The dental mesenchyme condenses and grows in response to the secondary signals sent by the epithelium. The more mesenchyme there is, the more the cervical loops grow laterally (and the less sharp the tooth becomes). This is implemented as a tendency for border cells to grow downwards in the *z* direction. That is counteracted by the amount of underlying mesenchyme:

$$\frac{\partial \mathbf{x}_{\text{border cell } i}}{\partial t} = \frac{d_{y_0}k_{\mathrm{Egr}}(1-d_i)}{\sqrt{(d_x^2 + d_y^2 + k_{\mathrm{Dgr}}^2)}} \tag{7}$$

$$\frac{\partial \mathbf{y}_{\text{border cell } i}}{\partial t} = \frac{d_{y_0}k_{\mathrm{Egr}}(1-d_i)}{\sqrt{(d_x^2 + d_y^2 + k_{\mathrm{Dgr}}^2)}} \tag{8}$$

$$\frac{\partial \mathbf{z}_{\text{border cell } i}}{\partial t} = \frac{d_{y_0}k_{\mathrm{Egr}}(1-d_i)}{\sqrt{(d_x^2 + d_y^2 + k_{\mathrm{Dgr}}^2)}} \tag{9}$$

$$d_x = d_{xo}c \tag{10}$$

$$d_y = d_{yo}c \tag{11}$$

$$c = 1 + \frac{k_{\mathrm{Mgr}}[\mathrm{Sec}]}{\sqrt{d_{xo}^2 + d_{yo}^2}} \tag{12}$$

Where d_{xo} and d_{yo} are the x and y components of equation (5) describing the growth of any kind of epithelial cell. k_{Mgr} is a model parameter that specifies the efficiency of Sec in promoting mesenchyme proliferation, and k_{Dgr} is a parameter describing the tendency to grow downwards when there is a minimal amount of mesenchyme.

Mesenchymal buoyancy. To model the resistance that the mesenchyme opposes to the invagination of the epithelium, each epithelial cell centre receives a displacement perpendicular to its apical cell surface (this assumes that the mesenchyme behaves like a compressed fluid):

$$\frac{\partial \mathbf{p}_i}{\partial t} = k_{\mathrm{Boy}}(1 - d_i)[\mathrm{Sec}]\hat{\mathbf{n}} \tag{13}$$

Where k_{Boy} is a model parameter describing how strong that force is, and $\hat{\mathbf{n}}$ is a unit vector normal to the epithelial cell surface. This vector is calculated as the sum of the vectorial product of all the vectors going from a cell to its neighbours (only the resulting vectors pointing apically are taken for each product).

Gene network parameters. Three growth factors are considered in the model, Act, Inh and Sec. Of these, Sec has different effect on epithelium and mesenchyme growth, and it could be modelled as two separate growth factors (Sec1 and Sec2) but because here we modulate all the other parameters related to Sec similarly, we consider Sec as a single node. The set of equations describing the change in the concentration of each growth factor in the immediate surroundings of an epithelial cell are:

$$\frac{\partial[\mathrm{Act}]}{\partial t} = \frac{k_{\mathrm{Act}}[\mathrm{Act}]}{1 + k_{\mathrm{Inh}}[\mathrm{Act}]} - k_{\mathrm{Deg}}[\mathrm{Act}] + k_{\mathrm{Da}}\nabla^2[\mathrm{Act}] \tag{14}$$

$$\frac{\partial[\mathrm{Inh}]}{\partial t} = -k_{\mathrm{Deg}}[\mathrm{Inh}] + k_{\mathrm{Di}}\nabla^2[\mathrm{Inh}] \tag{15}$$

$$\frac{\partial[\mathrm{Sec}]}{\partial t} = -k_{\mathrm{Deg}}[\mathrm{Sec}] + k_{\mathrm{Ds}}\nabla^2[\mathrm{Sec}] \tag{16}$$

The first equation states that all epithelial cells produce and secrete Act in an autoregulatory fashion, and the model parameter k_{Act} represents the strength of that autoregulation. The model parameter k_{Inh} describes the inhibitory effect of Inh over Act production. k_{Deg} is a generic degradation rate for all gene products. k_{Da}, k_{Di} and k_{Ds} are the diffusion constants for the Act, Inh and Sec molecular signals. The last term in each equation is the term describing the diffusion of each growth factor.

The same equations hold for mesenchymal cells except that the first term of the first equation and the two first terms of the fourth equation are omitted. This means that mesenchymal cells have no reaction terms, and thus only epithelial cells secrete the three growth factors. Epithelial cells become enamel knot cells once they reach the arbitrary Act concentration of one. This differentiation is irreversible. In that sense the signalling part of the model can be considered as an irreversible reaction-diffusion-like model. Enamel knot cells secrete Inh and Sec. Other cells secrete Inh after their differentiation state reaches the value of the model parameter k_{Int} (Int denotes initial threshold) and Sec after their differentiation state reaches the value of the model parameter k_{Set} (Set denotes secondary threshold). This reflects the stepwise upregulation of genes in the epithelium. The equations for the secretion of these signals are:

$$\frac{\partial[\mathrm{Inh}]}{\partial t} = [\mathrm{Act}] - k_{\mathrm{Deg}}[\mathrm{Inh}] + k_{\mathrm{Di}}\nabla^2[\mathrm{Inh}] \tag{17}$$

$$\frac{\partial[\mathrm{Sec}]}{\partial t} = k_{\mathrm{Sec}} - k_{\mathrm{Deg}}[\mathrm{Sec}] + k_{\mathrm{Ds}}\nabla^2[\mathrm{Sec}] \tag{18}$$

Where k_{Sec} is a model parameter describing Sec secretion. Similar results are obtained if Sec secretion is dependent on Act concentration.

Initial conditions. The initial conditions consist of an epithelial layer representing the tip of the oral epithelium invagination and 20 underlying mesenchymal cell layers. Each cell is a three-dimensional volume including the cell itself and its immediate extracellular surrounding space. Initially each cell has a columnar hexagonal morphology. Epithelial and mesenchymal cell layers are composed of seven cells for the seal simulations. The cells in each layer are arranged in a hexagon with each cell at an initial distance ($||p_o||$) of one to each of its neighbours (see Supplementary Fig. 1). Initially each epithelial cell has six epithelial neighbours, except for cells on the borders that have three or four depending on the location (see Supplementary Fig. 1). Each epithelial cell has one underlying mesenchymal neighbour. During the simulation an external constant source of Act exists next to epithelial cells in the border of the tooth as found experimentally[19,39]. Other than that all initial gene product concentrations and differentiation are set to zero in all cells. All parameters are dimensionless.

31. Salazar-Ciudad, I., Garcia-Fernández, J. & Solé, R. V. Gene networks capable of pattern formation: from induction to reaction-diffusion. *J. Theor. Biol.* **205,** 587–603 (2000).
32. Jaeger, J. *et al.* Dynamic control of positional information in the early *Drosophila* embryo. *Nature* **430,** 368–371 (2004).
33. Newman, T. J. Modeling multi-cellular systems using sub-cellular elements. *Math. Biosci. Eng.* **2,** 611–622 (2005).
34. Honda, H., Motosugi, N., Nagai, T., Tanemura, M. & Hiiragi, T. Computer simulation of emerging asymmetry in the mouse blastocyst. *Development* **135,** 1407–1414 (2008).
35. Graner, F. & Glazier, J. A. Simulation of biological cell sorting using a two dimensional extended Potts model. *Phys. Rev. Lett.* **69,** 2013–2016 (1992).
36. Hammer, Ø. & Bucher, H. Models for the morphogenesis of the molluscan shell. *Lethaia* **38,** 111–122 (2005).
37. Butler, P. M. The ontogeny of molar pattern. *Biol. Rev. Camb. Philos. Soc.* **31,** 30–69 (1956).
38. Jernvall, J. Mammalian molar cusp patterns: developmental mechanisms of diversity. *Acta Zool. Fenn.* **198,** 1–61 (1995).
39. Åberg, T., Wozney, J. & Thesleff, I. Expression patterns of bone morphogenetic proteins (Bmps) in the developing mouse tooth suggest roles in morphogenesis and cell differentiation. *Dev. Dyn.* **210,** 383–396 (1997).