Developmental Nonlinearity Drives Phenotypic Robustness

Rebecca M. Green  
*University of Calgary*

Jennifer L. Fish  
*University of Massachusetts Lowell*

Nathan M. Young  
*University of California San Francisco*

Francis J. Smith  
*University of Colorado Anschutz Medical Campus*

Benjamin Roberts  
*University of Massachusetts Lowell*

*See next page for additional authors*

Follow this and additional works at: [https://ecommons.luc.edu/biology_facpubs](https://ecommons.luc.edu/biology_facpubs)

Part of the Biology Commons

**Recommended Citation**

Green, Rebecca M.; Fish, Jennifer L.; Young, Nathan M.; Smith, Francis J.; Roberts, Benjamin; Dolan, Katie; Choi, Irene; Leach, Courtney L.; Gordon, Paul; Cheverud, James M.; Roseman, Charles C.; Williams, Trevor J.; Marcucio, Ralph S.; and Hallgrimsson, Benedikt. Developmental Nonlinearity Drives Phenotypic Robustness. *Nature Communications*, 8, 1970: 1-12, 2017. Retrieved from Loyola eCommons, Biology: Faculty Publications and Other Works, [http://dx.doi.org/10.1038/s41467-017-02037-7](http://dx.doi.org/10.1038/s41467-017-02037-7)

This Article is brought to you for free and open access by the Faculty Publications and Other Works by Department at Loyola eCommons. It has been accepted for inclusion in Biology: Faculty Publications and Other Works by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.

This work is licensed under a [Creative Commons Attribution 4.0 International License](http://creativecommons.org/licenses/by/4.0/).
Authors
Rebecca M. Green, Jennifer L. Fish, Nathan M. Young, Francis J. Smith, Benjamin Roberts, Katie Dolan, Irene Choi, Courtney L. Leach, Paul Gordon, James M. Cheverud, Charles C. Roseman, Trevor J. Williams, Ralph S. Marcucio, and Benedikt Hallgrimsson

This article is available at Loyola eCommons: https://ecommons.luc.edu/biology_facpubs/84
Developmental nonlinearity drives phenotypic robustness

Rebecca M. Green, Jennifer L. Fish, Nathan M. Young, Francis J. Smith, Benjamin Roberts, Katie Dolan, Irene Choi, Courtney L. Leach, Paul Gordon, James M. Cheverud, Charles C. Roseman, Trevor J. Williams, Ralph S. Marcucio & Benedikt Hallgrímsson

Robustness to perturbation is a fundamental feature of complex organisms. Mutations are the raw material for evolution, yet robustness to their effects is required for species survival. The mechanisms that produce robustness are poorly understood. Nonlinearities are a ubiquitous feature of development that may link variation in development to phenotypic robustness. Here, we manipulate the gene dosage of a signaling molecule, Fgf8, a critical regulator of vertebrate development. We demonstrate that variation in Fgf8 expression has a nonlinear relationship to phenotypic variation, predicting levels of robustness among genotypes. Differences in robustness are not due to gene expression variance or dysregulation, but emerge from the nonlinearity of the genotype–phenotype curve. In this instance, embedded features of development explain robustness differences. How such features vary in natural populations and relate to genetic variation are key questions for unraveling the origin and evolvability of this feature of organismal development.
Waddington proposed that selection tends to stabilize development along particular paths, a phenomenon he called "canalization". He tested this idea by selecting for an induced trait in the presence of a teratogen (e.g., ether and the bithorax phenotype) and obtained individuals in which the trait appeared without the teratogen. He hypothesized that selection had stabilized development around the induced trait such that it no longer needed the environmental stimulus. Concurrent work by Waddington and others showed that mutations with major effects tended to be more variable than the wild type. This observation was also explained by invoking canalization. Mutations were hypothesized to increase variance by disrupting evolved mechanisms that buffered variation around a phenotypic mean. This tendency for resistance to perturbation in development, or robustness, is widely thought to be a fundamental property of complex life. Yet, the mechanisms responsible for promoting and modulating robustness remain largely unknown.

Wagner et al. defined canalization as suppression of phenotypic variation among individuals due to insensitivity to either genetic or environmental effects. This definition hinges on a distinction between the frequency distribution of the genetic or environmental factors that cause variation and the magnitudes of phenotypic effect associated with those factors. A mutation or environmental effect disrupts or decreases canalization when phenotypic variance is increased, while all other genetic or environmental effects are unchanged.

Two kinds of mechanisms have been proposed to explain canalization. In one, specific molecular mechanisms such as heat shock and other chaperone proteins, or microRNAs buffer against perturbations and suppress the expression of variation. In the other, canalization emerges from redundancies, feedback loops, and other features of developmental systems. These explanations are not mutually exclusive, and multiple mechanisms may act simultaneously at different levels of development. However, they differ in that one posits the existence of organism-wide buffering processes that reduce variation, while the other holds that robustness emerges from the same mechanisms that generate variation in specific traits. A common feature of developmental systems explanations for robustness is the importance of nonlinearity. Within tissues, processes such as the diffusion of a morphogen are nonlinear in ways that depend on anatomical context. Genetic variation influences the phenotype via developmental processes that act at different scales, times, and locations within the organism, complicating the relationship between genotype and phenotype. Therefore, it is not clear how nonlinearities in specific mechanisms translate to quantitative relationships between genetic and phenotypic variation.

Lewontin introduced the genotype–phenotype (G–P) map to conceptualize relationships between genetic and phenotypic variation. G–P maps are often nonlinear, as evident in dominance and epistasis. While much has been learned about the developmental mechanisms that construct vertebrate morphology, much less is known about the relationship between developmental and quantitative phenotypic variation. Alberch conceptualized how nonlinearities in specific features of developmental systems translate to quantitative relationships between developmental processes and phenotypic variation. In the developmental landscape, nonlinear relationships between developmental processes and phenotypic variation can arise from the interaction of genetic and environmental factors.
We previously demonstrated significant nonlinearity in the relationship between sonic hedgehog signaling and embryonic facial shape. Variation in the three-dimensional morphology of the face is far removed from nonlinear molecular processes or the theoretical dynamics of gene regulatory networks. For this reason, it is not at all clear that the theoretical predictions that link nonlinearity to phenotypic variance should hold across the vast complexity of the G–P relationship. To test the hypothesis that a nonlinear G–P relationship predicts variation in robustness, we examine how variation in Fgf8 expression affects the mean and phenotypic variance for craniofacial shape.

Fgf8 is appropriate for this study as it drives a central pathway in craniofacial development. Fgf8 is a signaling factor that is expressed in the facial and oral ectoderm, where it directs craniofacial pattern and polarity. Fgf8 is absolutely required for proper development of facial structures. Fgf8-expressing cells form a boundary with Shh-expressing cells to form the frontonasal ectodermal zone, which directs the outgrowth of the facial prominences and has also been implicated in their evolution.

We predict that Fgf8 expression relates nonlinearly to craniofacial phenotype. Further, we predict that the shape of the curve relating mean phenotype to Fgf8 level will dictate the phenotypic variance within and between genotypes. Genotypes falling on the steeper portions of the curve will have higher variances (differences among individuals within genotype) than the genotypes falling on flatter portions. Likewise, different genotypes that fall along the steeper portions of the curve will have higher genetic variances, while those along the flatter portion of the curve will show little phenotypic variation.
of gradations in $Fgf8$ dosage (Fig. 2c, Supplementary Table 1). Mean $Fgf8$ levels in the head for the nine genotypes relative to the wild-type embryos from the $Fgf8$neo series vary significantly (ANOVA, df = 69, 8, $P < 1 \times 10^{-26}$), ranging from 0.14 to 1.1 (Fig. 2c), yet we detect no difference in the variance of gene expression across the genotypes (Levene’s test, df = 69, 8, $P = 0.2043$). Further, by deleting $Fgf8$ in two different ways, we are able to show consistency between different mechanisms of $Fgf8$ loss. Facial phenotype is assessed by geometric morphometric analysis48,50. Here, the perturbation is the generation of a genotype—phenotype map. To determine the relationship between $Fgf8$ expression and phenotypic variation, we used Morrissey’s34 quantitative model for nonlinear G–P maps. This model produces a prediction of the amount of variance that should be observed given a nonlinear G–P map. To generate the curve used to test Morrissey’s model, we fit the $Fgf8$ gene expression data, and the phenotypic data (3D landmark data) to a von Bertalanffy curve using least-squares regression. The phenotype data used was the regression score from a multivariate regression of our normalized Procrustes coordinates on $Fgf8$ level—which generates single variable shape score48. These curves are shown in Fig. 4a, b.

**Fig. 3** Shape changes in response to decreased $Fgf8$ gene dosage. Principal component analysis (PCA) of shape at E10.5 (a, b) and P0 (c, d). Gray embryos (E10.5) show shape change trajectories for PC1 (horizontal) and PC2 (vertical), and the middle vertical image represents the mean shape for each time point. Gold skulls show the same shape change trajectories for the P0 data. a, c PC plots colored by $Fgf8$ level with warm colors representing wild-type embryos and cool colors and purples showing around 20% mRNA expression (mean per group by qRT-PCR). b, d Coloration by genotypic series, genotypes separate by allelic series, but the differences between low PC1 and low PC2 are small. The mean shape of individual genotypes is already significantly different (Procrustes permutation test, $P < 0.001$) by E10.5. d Coloration by genotype as used in the rest of the paper. A total of 187 neonates were analyzed and divided between groups as follows: WT (+/+), Flox/+, Flox/−, Neo/+, Neo/−, Crect. A similar pattern is seen at P0, showing that the correlation is preserved throughout embryogenesis (Fig. 3c, d).

**Loss of $Fgf8$ affects phenotypic variance.** The Morrissey model, based on the mean and standard deviation of our $Fgf8$ gene expression data, predicts that variation in $Fgf8$ expression has little effect on shape metrics (phenotypic value or regression score) when $Fgf8$ expression is >40% of the wild-type level, while below this point, variation in $Fgf8$ expression produces increasingly large effects on the mean phenotype. Figure 4c, d shows the predicted relationship between the variance of $Fgf8$
expression and phenotypic variance at four mean expression levels for the E10.5 and P0 samples. These results show that the variance of Fgf8 expression will have little effect on phenotypic variance when Fgf8 level is >50% of the wild-type level, while the phenotypic variance becomes increasingly sensitive to gene expression variance below this point.

Figure 5a, b shows the individual-level data for the regression of shape against mean Fgf8 level. The von Bertalanffy curve explains 54% of the phenotypic variance at E10.5 and 84% at P0. Fgf8 expression measured by RT-PCR in the head relates nonlinearly to craniofacial morphology. Following the prediction of the Morrissey model, when Fgf8 levels are above 40% of wild-type levels, the effect on mean shape is minimal. Below this point, however, the phenotype deviates sharply. When Fgf8 expression levels are reduced below 40% of wild-type levels, small differences in Fgf8 expression have large phenotypic effects.

To determine whether nonlinearity predicts robustness, we plotted variance in face shape against Fgf8 expression across genotypes. No change in shape variance, measured as the Procrustes variance or morphological disparity48,49, is seen until Fgf8 expression drops below 40% of wild-type levels (Fig. 5c). As predicted, shape variance dramatically increases below 40% expression in both the E10.5 and P0 samples, corresponding to the point at which the phenotype becomes sensitive to Fgf8 levels (P values between groups—Supplementary Table 2). The only exception is for E10.5 Fgf8; Crect embryos, likely due to the fact that Creact does not activate until E9.5. By P0, this group has significantly increased phenotypic variance (Supplementary Table 2). At P0, the Fgf8neo/− embryos are so highly dysmorphic that most of them could not be landmarked and so were not included in the variance analysis (Supplementary Fig. 1).

Effects on gene expression. To eliminate the possibility that differences in genetic variance across the allelic series account for the differences in phenotypic variance, we quantified genetic variance from high-resolution SNP data. These results show no relationship between genetic variance and phenotypic variance across the allelic series (Supplementary Fig. 2).

We determined the genome-wide changes in expression across the allelic series at E10.5 using RNAseq. We reduced the transcriptome data using PCA. The pattern of gene expression within genotypes varies across the allelic series. PCI of the transcriptome accounts for 44% of variation in gene expression. We interpret this PC to reflect the coordinated genome-wide changes in gene expression across the allelic series. This PCI ordinates the allelic series (Fig. 6a). Mean Fgf8 expression level by genotype accounts for 30% of the variation in PCI of the transcriptome (Fig. 6b). For further analysis, the data were separated into three groups: all genes, the MapK Kegg pathway, and a hand-curated list of 15 known, direct Fgf target genes (Supplementary Table 3). The MapK Kegg pathway was selected as Fgf signaling falls within the MapK signaling cascade. In all the genes and in the MapK pathway, there appears to be a curve in the data; however, each group is statistically different from its neighbor (Fig. 6c, d). This nonlinearity becomes more pronounced for the Fgf downstream targets. For these genes, expression level does not differ significantly among the heterozygote genotypes (T test, P = 0.96; Fig. 6e). The lack of change between these groups generates a flat region in the curve with an inflection point at 40–50% of the wild-type Fgf8 level, demonstrating nonlinearity.

To test the hypothesis that the low Fgf8 expression genotypes have increased phenotypic variation because of less coordinated
Among individuals. High correlations indicate a high degree of consistency among individuals within genotypes. We performed this analysis for both genome-wide and for each of the two groups of genes known to be downstream of Fgf8. This analysis revealed no evidence of dysregulation of gene expression across the allelic series. The pairwise correlations genome wide or within likely downstream targets show no detectable pattern across genotypes (Fig. 6f–h).

To determine whether the transcriptomic data show evidence of compensatory changes that could explain the lack of phenotypic response above 40% of the wild-type Fgf8 expression level, we searched for significant correlations between groups of genes and Fgf8 level across individuals and genotypes. Resampling revealed elevated ($P < 0.05$) correlations for the reactome Fgf downstream-signaling pathway, but not for Wnt, apoptosis, MapK Kegg, and hedgehog pathways. Eight individual genes fell outside of the 95% confidence interval based on genome-wide resampling of a similar number of genes. This list includes Fgf4 and Trib3 that are negatively correlated with Fgf8, and Fgf17, Etv4, Prkg, Spry1, Spry4, and Rictor that correlate positively. The two Sprouty (Spry) genes and Etv4 are known to be downstream of Fgf8, suggesting that Fgf8 signaling modulates downstream genes across the entire range of expression. A MANOVA shows that the genotypes vary significantly in the expression of these downstream genes (Pillai’s trace $= 2.1$, $P < 1 \times 10^{-3}$). As a confirmation, we performed RT-PCR analysis on these eight genes and compared them against the Fgf8 levels for each sample (Fig. 7). While Fgf4 failed to reach statistical significance, Trib3 does appear to be weakly, but significantly negatively correlated with Fgf8 level. All other genes were positively and significantly correlated with Fgf8 level. Fgf17, for example, trends toward mild upregulation in the Neo/+ group (1.27 $\pm$ 0.16 vs. 1.06 $\pm$ 0.32, Student’s $t$ test, $P = 0.15$). Mean levels by genotype and standard deviations have been listed in Supplementary Table 4. These results suggest that there may be genes that change in expression to compensate for loss of Fgf8, though this requires further investigation.

**Discussion**

We show that nonlinearity in the G–P relationship for Fgf8 expression predicts phenotypic robustness. Progressive reduction in Fgf8 yields a nonlinear relationship to phenotype, affecting both mean facial shape and the magnitude of phenotypic variance. Development tolerates a large amount of change in Fgf8 expression around wild type, but only to a point, after which small changes in Fgf8 lead to large changes in phenotype, thus permitting more morphological variance to be generated in a population for a given amount of variation in Fgf8. These findings show that nonlinearity in a single pathway can propagate across the many levels of organization (molecular, cellular, tissue, etc.) that channel information from genotype to phenotype, providing a viable mechanistic explanation for canalization (Fig. 1).

Our results are consistent with the hypothesis that the nonlinear G–P map for Fgf8 explains the differences in phenotypic variance across the Fgf8 allelic series. There are minor differences among individuals in Fgf8 expression within each genotype for the allelic series. Our model predicts that these minor differences will translate to different magnitudes of phenotypic variance along the curve that describes the relationship between Fgf8 expression and the mean cranial phenotype at each point along the curve. This result implies that robustness can emerge in developmental context as a consequence of nonlinearities in development. This contrasts with explanations for canalization that invoke dedicated mechanisms such as heat shock proteins that regulate variability organism wide. Our results do not
preclude the existence of such mechanisms, but they provide an additional and, perhaps, more general explanation for genetic and environmental influences on phenotypic robustness.

An alternative explanation to the changes in variance along the range of Fgf8 expression is that disruptions to Fgf8 expression dysregulate downstream gene regulatory networks, producing increased variance in gene expression that translates to increased phenotypic variance. Such disruptions might be specific to downstream targets of Fgf8 or be more widespread. By this explanation, differences among individuals are greater at the lower range of Fgf8 expression because these individuals also vary in expression of downstream genes. It predicts that as Fgf8 dosage falls below the threshold, the variance of downstream gene expression increases. This explanation relies on the idea that extreme changes in gene expression may have systemic destabilizing effects on development. This is implicit in the Hsp90 explanation for the source of robustness, as well as in several older explanations for canalization such as Lerner’s genetic homeostasis model. However, we found no evidence of increased variance of gene expression, suggesting that the increased phenotypic variance in genotypes producing low levels of Fgf8 is not attributable to greater instability of the downstream gene regulatory network.

**Fig. 6** Gene expression changes across the Fgf8 allelic series. a PC 1 and 2 plot of RNAseq data (18 samples). No differences in dispersion are observed between groups. b Relationship between PC1 of the RNAseq data and Fgf8 level for each sample as quantified by RT-PCR. The blue line shows the line of best fit, gray shows the 20% error around the line. c–e Average absolute value fold change (dot) and average absolute value standard error of the fold change (error bar) between each mutant genotype and wild type, c all measured genes, d MapK Kegg gene list (174 genes), and e Fgf8 downstream targets (15 genes). The asterisk represents $P < 0.05$ (bootstrap resampling) between nearest-neighbor groups (shown between the groups). The white vertical line shows an Fgf8 level of $\sim 40\%$. f–h Z-transformed covariance between each embryo within a genotype on f all genes, g MapK Kegg gene list (174 genes), and h 15 known Fgf8 downstream targets (Supplementary Table 3). The asterisk represents $P < 0.05$ (bootstrap resampling) between group and wild type.
We did find, however, that genes downstream of Fgf8 respond nonlinearly to Fgf8 expression. In other words, the increased phenotypic effects at low Fgf8 levels are mirrored by increased changes in gene expression, particularly in genes known to be downstream of Fgf8. This suggests that the nonlinear G–P map is a feature of a larger gene regulatory pathway, and that the phenotypic effects at low Fgf8 levels are occurring because many genes are more responsive to Fgf8 levels within that range than at levels closer to the wild type.

Interestingly, the phenotypic effects of the loss of Fgf8 become more marked between E10.5 and P0. At P0, the genotypes separate more clearly and the increase in phenotypic variance at the steep part of the curve becomes more marked. Fgf8 is expressed throughout facial prominence outgrowth and face formation53. Our results suggest that the effects of perturbing Fgf8 expression below the threshold of 40% are exacerbated during late embryonic and fetal development.

Fig. 7 RT-PCR validation of correlation eight genes with Fgf8 level. Thirty-seven samples from across the genotypes were analyzed for each of the eight genes plus Fgf8 and modeled for a linear relationship. The linear relationship is shown in red (line±SE shaded). R2 values and the adjusted P values from the linear model are shown.
Here, we build on earlier work in which we show a nonlinear G–P map for Shh expression and facial shape in chicks. This study did not determine how phenotypic variance is modulated along the expression curve, however. Further, that study manipulated Shh expression directly rather than via a genetic model as we have done here. The advantage of the genetic approach is that we can eliminate experimental error as a source of among-individual variance within groups.

Our findings have important implications for the evolvability of morphology. Applying Morrisey’s model shows that even with strong selection on midfacial shape and substantial expression variation in Flg8 levels, there would be little to no response to selection on facial morphology through alterations of Flg8 expression levels. The correlated response of Flg8 expression would be very low. On the other hand, at lower mean Flg8 expression levels, response to selection on midfacial shape would be achieved, at least in part, by changes in Flg8 expression. There would be a substantial correlated response in Flg8 levels. These contrasting results flow directly from the nonlinear relationship between Flg8 levels and midfacial morphology and suggest that while Flg8 clearly plays a pivotal role in craniofacial development, it is unlikely to contribute directly to microevolutionary changes in craniofacial form under a wide range of expression levels, from 40% of wild-type expression to full wild-type expression.

Similarly, the nonlinear G–P map for Flg8 expression and craniofacial shape helps us understand a puzzling and emerging trend in the genetics of complex traits. Why is it that the genes controlling craniofacial shape help us understand a puzzling and emerging trend in craniofacial form under a wide range of expression levels, from 40% of wild-type expression to full wild-type expression?

But how do nonlinear G–P maps for key developmental factors such as Flg8 arise in the first place? Developmental origins of nonlinearities contribute to the variation in the genetics of complex traits. Why is it that the genes known from developmental biology to play major roles in the construction of morphology so often appear to play minimal roles in determining the variation of that morphology? Studies of craniofacial shape variation in mice and humans reveal a growing list of causal variants. While some have known roles in facial development, many of the major players such as Shh or Flg8 are conspicuously absent from these lists. Nonlinear G–P maps for such central genes would explain this result.

For embryos, pregnant dams were sacrificed at E0.5. Embryos were dissected on ice and fixed in glutaraldehyde prior to CT scanning. Neonates were sacrificed at embryonic day (E) 10.5 based on visualization of a postcoital plug at E0.5. Embryos were dissected on ice and fixed in 4% paraformaldehyde and 5% glutaraldehyde prior to µCT scanning. Neonates were killed in CO2 on ice and then fixed in 4% paraformaldehyde and 5% glutaraldehyde prior to µCT scanning.

Mouse experiments were approved by the UC Denver Institutional Animal Care and Use Committee (CreCre mice) and by the UCSC and University of Massachusetts Lowell Institutional Animal Care and Use Committee (Neo series mice).

### Methods

**Mouse breeding and embryo generation.** The Fgf8flox series is a five-member series generated from a combination of the neomycin insertion into the intron between exons 2 and 3 of the Fgf8 locus and a null allele generated from loss of exon 2. The Fgf8Cre series contains combinations of a floxed allele for Fgf8, a null allele for Fgf8, and then Fgf8 is deleted from the ectoderm around E9.5 using an ectodermal Cre (CRECT) (Fig. 2). The two series of mice were generated independently by different labs (Cret, T. Williams) (Nes, R. MarcziaJ. Fish). Both series of Fgf8 mice were generated from the Fgf8flpflxed allele originally developed by Meyers et al. The neo cassette was maintained in the Fgf8Neo mice. To generate the floxed allele for the CRET series, the neomycin resistance cassette was removed by crossing these mice to β-actin–flp (B6.Cg-Tg(ActFLP) Pe 90.1Cre(CAG-FLPe)/123M.L). Deletion constructs were developed by crossing with β-actin–Cre (EvB2N-Tg(ActB-cre2)Mt/2Fr)'), to delete exons 2 and 3 from all cells.

To generate the Fgf8flcox series, crosses were performed between mice that were heterozygous for the Neo (flp) allele or heterozygous for the Neo (flp) allele and the null allele. The Neo allele was genotyped with the following primers (5′–3′): F: CTG CAC AAG GCC AGA GA G; R: AGC TCC CGC TGG ATT CCT C. The null allele (UCSF/Umass) was genotyped with the following primers (5′–3′): F: GCC GTG TGA ATT TGA CCT GAG GCC; R: GAA AAC GCC ATC GCA GGC TTC TGC. The null and Neo alleles can be genotyped simultaneously at an annealing temperature of 58 °C. The floxed allele was genotyped using the following primers: (5′–3′) EM 99: CTT AGG GCT ATC CAA CCC ACC ATC and EM32: GGT CTC CAC AAT GAG CTT C. The null allele (UCDenver) was genotyped using EM41: AGC TCC CGC TGG ATT CCT C and EM99. These three can also be genotyped simultaneously at 58 °C. The Cre deletion series was generated by crossing the Cret homozygous Cre (Fig. 3), with the null allele, and then males from this cross were crossed to Fgf8floxplox females on an FVB background. Genotyping for this allele was performed using general Cre primers (5′–3′): Cre1: GCT GGT TAG CAC GGC AGG TGT AGA G; Cre3: GGC CAT CTT CCA GCA GGC GCA CC with a 67 °C annealing temperature.

**SNP analysis to estimate genetic variance.** Following genotyping, five DNA samples per each wild-type (WT), Neo+/+, WT/−, Neo/Neo, and Neo−/− were sent to GeneSeek Inc. (a Neogene company, Lincoln, Nebraska USA). DNA samples were run on the GigaMaga mouse genotyping chips (Ilumina), for a total of 144,230 (200,000) SNPs. After quality control, oligo and removal of low-frequency SNPs, we performed analyses using 133,559 SNPs on each of 17 samples, 4–5 per group. QC and SNP calls were done using the GenomeStudio Package (Ilumina) by...
Scanning and landmarking. All samples were µCT scanned on a µCT35 scanner to visualize facial shape. Prior to scanning, embryos were submersed in CytoCon Ray II (iothalamate meglumine) contrast agent for 1 h, and then scanned at 7.5-µm resolution. Neones were scanned at 19-µm resolution without contrast agent to allow resolution of the bone. Scans were then reconstructed and landmarked using Meshlab (Version 1.3.2, Visual Computing Lab, meshlab.sourceforge.net) (embryos) or Amira (Version 5.2, FEI) (neonates). Landmarks for embryos were as developed by Percival et al.50. Neonate landmarks for the cranium are from Gonzalez et al.13, with the addition of the landmarks on the mandible. Landmarks for each age group were placed by a single observer who was blinded to a genotype. A total of 38 landmarks were placed on the embryos and 76 were placed on the neonates. Samples with shrinkage artifacts, or missing landmarks were removed from analysis.

Geometric morphometrics. Landmark data were imported into R, and Procrustes superimposition was performed to remove scaling and orientation differences between samples using the Geomorph package72,73 in R74. Embryo data were regressed against tail somite number to remove ontogenetic effects before further analysis. Neonate data were regressed against centroid size only. Background effects due to lab of origin were mitigated by removing the difference between the wild-type groups from all specimens. A total of 187 neonates were analyzed and divided between groups as follows: WT (+/+): 22, Fllox/+; Crect +/−15, Fllox/+; Crect −/−15, Fllox/+; Crect −/+16, Fllox/+; Crect −/−16, Fllox/−; Crect +/−19, Neo/Ne = 12, Neo/Ne = −3 (all landmarks present). A total of 156 embryos were analyzed and divided between groups as follows: WT (+/+): 27, Fllox/+ = 15, Neo/Ne = +30, Fllox/+ = ±16, Fllox/+; Crect = 16, Fllox/+; Crect −/−19, Neo/Ne = 12, Neo/Ne = −8. Sample sizes were based on previous work10. Our power analysis shows that 10 embryos are needed to detect a 15–30% increase in variance and five embryos are needed to detect a 20–30% increase in variance with a power of 0.8. Due to the large number of genotypes, we focus on trends across the data set rather than between-group differences. The size- and lab-normalized shapes (Procrustes coordinates) were then regressed against Fgf8 level in Figs. 4 and 5. Residuals from both the age regression and the Fgf8 regression were obtained using a linear model, as implanted by the procD.Allometry function in geometric morphometrics. To represent these regressions as a single variable, we used the common allometric coefficient (CAC). When calculated from a pooled analysis with multiple groups, this is mathematically identical to a regression score 2 and plots these values as the dependent variables against the independent variables (Fgf8 level and tail somite stage).

Modeling of phenotypic variance. To model the relationship between Fgf8 expression and the phenotypic mean and within-genotype variance, we used Morrissey’s model for the quantitative genetics of nonlinear G–P maps34. This model shows how the phenotypic mean is determined by the functional relationship between developmental processes (e) and the phenotype (z): 

$$\tau = f(e)N(e, \sigma_1, \sigma_2^2)de,$$

(1)

where f(e) is the functional relationship between the developmental process and the phenotype and N(e, \sigma_1, \sigma_2^2) is the normal distribution of developmental values (Fgf8 expression) with the specified mean and variance. The relationship between developmental and phenotypic variance is given by 

$$\sigma_2^2 = \Phi \sigma_1^2,$$

(2)

where 

$$\Phi = \int f'(e)p(e)de,$$

(3)

f'(e) is the first derivative of function f(e) and p(e) is the frequency of specific developmental values.

Modeling of the G–P curve. We fit the phenotype to Fgf8 expression at E10.5 and at P0 using a nonlinear least-squares regression to a von Bertalanffy curve of the formula:

$$z = L_m - (L_m - L_0) e^{-k}$$

(4)

where L_m is the maximum phenotype, L_0 is the mean phenotype at zero expression (y-intercept), and k is a rate constant describing the decrease in slope per unit of e. In this curve, the initial rate of change of a phenotype given e decreases at a rate proportional to k until it reaches an asymptote (L_m).

RNA collection for gene expression analyses. E10.5 embryos were dissected into PBS on ice and snap frozen at −80 °C. Heads were dissected from between the mandibular arch and the hyoid arch. All RNA work was performed on the RNA extracted from the head. RNA was extracted in batch preps using Trizol. cDNA was made from 500 ng of RNA in a 20-µl reaction mix using an iScript cDNA synthesis kit (Bio-Rad).

qPCR. Reverse transcription quantitative real-time PCR (qRT-qPCR) was performed as previously described77. Briefly, we use a C1000 Thermal Cycler with a CFX96 Real-Time System (Bio-Rad). Forward and reverse primers, 2 µl of cDNA, RNAse-free dH2O, and SYBR–Select Master Mix (ThermoFisher), containing dNTPs, iTaq DNA polymerase, MgCl2, SYBR Green I, enhancers, stabilizers, and fluoroscein, were manually mixed in a 20-µl reaction to amplify each cDNA of interest. Primer sequences were: GAPDH (F: 5’-AGGTGCTGTGTAGGAATTG-3’; R: 5’-GGGCTGTGTAGGGCAAA-3’); and FGF8 (F: 5’-GATTGTTTCCGCAGGATGCAG-3’; R: 5’-GACAGGTCTCCTACATGGCATG-3’). Each sample was run in triplicate, all results were normalized to the expression of GAPDH, and fold changes were calculated using the delta–delta Ct method38. Primers for qRT-PCR were selected for optimal G/C concentrations and tested for ideal melt curves and optimized for amplification efficiency: GAPDH, 92% at 61.5 °C and FGF8, 102% at 61.6 °C.77. Primers for Fgf8 were located in the 5’ end of the transcript to prevent detection of nonfunctional transcript generated from the Neo or LacI insertions.

Real-time PCR quantification of the RNAseq data was performed as follows. cDNA was generated using the Maxima First Strand Kit (Thermo-Fisher) and amplified using the IDT mastermix and IDT PrimeTime probes and primers (Mm.

PT.58.10694850, Mm.PT.58.7996582, Mm.PT.58.49383184, Mm.PT.58.29112396, Mm.PT.58.32292921, Mm.PT.58.43808967, Mm.PT.58.42037828 g). Samples were run on an Applied BioSystems QunatiStudio 6. Data were normalized by averaging Gapdh and β-actin expression levels. deltaCT values were used in all downstream analysis. Correlation analysis was performed in R. The mean deltaCT for the controls was calculated before the log transformation for each sample, resulting in a slight alteration of the wild-type mean from 1.

RNAseq. RNA quality was assessed using an Agilent TapeStation and RNA integrity number (RIN) were selected for optimal G/C concentrations and tested for ideal melt curves and optimized for amplification efficiency: GAPDH, 92% at 61.5 °C and FGF8, 102% at 61.6 °C.77. Primers for Fgf8 were located in the 5’ end of the transcript to prevent detection of nonfunctional transcript generated from the Neo or LacI insertions.

Real-time PCR quantification of the RNAseq data was performed as follows. cDNA was generated using the Maxima First Strand Kit (Thermo-Fisher) and amplified using the IDT mastermix and IDT PrimeTime probes and primers (Mm.

PT.58.10694850, Mm.PT.58.7996582, Mm.PT.58.49383184, Mm.PT.58.29112396, Mm.PT.58.32292921, Mm.PT.58.43808967, Mm.PT.58.42037828 g). Samples were run on an Applied BioSystems QunatiStudio 6. Data were normalized by averaging Gapdh and β-actin expression levels. deltaCT values were used in all downstream analysis. Correlation analysis was performed in R. The mean deltaCT for the controls was calculated before the log transformation for each sample, resulting in a slight alteration of the wild-type mean from 1.

Received: 30 October 2016 Accepted: 2 November 2017

Published online: 06 December 2017

References

Acknowledgements
This work was supported by grants NIH R01 2R01DE019638 to R.S.M. and B.H., NSERC 238992-17 to B.H. and C.C.R., and NIDCR R01 DE019843 to T.J.W. We thank Richard Hawkes for his valuable comments on the manuscript.

Author contributions
R.M.G., J.L.F., B.H., R.S.M. and T.J.W. designed the experiments. R.M.G., J.L.F., I.C. and K.D. generated the embryos for analysis. R.M.G. and F.J.S. did the microCT scanning and landmarked the embryos. R.M.G. and B.H. analyzed the morphometric data. B.R. and K.D. generated the RNA, DNA, and ran the qPCR along with C.L.L., J.L.F. and R.M.G. analyzed the qPCR data. R.M.G., C.C.R. and P.G. analyzed the RNAseq data. C.C.R. analyzed the S.N.P. data. N.M.Y., J.M.C., C.C.R., R.S.M., B.H., J.L.F. and R.M.G. helped interpret the data and develop the initial model. J.M.C. generated the mathematical modeling. R.M.G., J.L.F. and B.H. wrote the paper. All authors revised and approved the final manuscript.

Additional information
Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-017-02037-7.

Competing interests: The authors declare no competing financial interests.

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

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2017