Supporting Information

Mallarino et al. 10.1073/pnas.1011480108

SI Materials and Methods

In Situ Hybridizations and Immunohistochemistry. For immunohistochemistry, sections were blocked with 3% BSA in PBS containing 0.1% Triton-X 100 for 1 h, incubated overnight with primary antibody at 4 °C, washed in PBS, incubated for 1 h with secondary antibody, and washed with PBS. Immunostaining was performed using anti-*TGFβIIr* (sc-400; Santa Cruz), anti-*TGFβ1*, *-β2*, *-β3* (sc-146, sc-90, sc-82, respectively; Santa Cruz), anti–*β-catenin* (610153; BD Transduction Laboratories), and anti-*Dkk3* (kindly provided by Dr. Christof Niehrs, German Research Cancer Center, Division of Molecular Embryology, Heidelberg, Germany). Antibodies were used at dilutions of 1:50–1:200. Reactions were visualized with Alexa dye conjugated secondary antibodies (Molecular Probes) at 1:500 dilution in 3% BSA/PBS/Triton-X 100. For controls, sections were incubated with PBS instead of primary antibodies but no specific cellular staining was observed.

Alkaline Phosphatase. Embryos were blocked with 3% BSA in PBS containing 0.1% Triton-X 100 for 1 h and incubated with an AP-conjugated secondary antibody (Jackson ImmunoResearch). The signal was detected using a combination of nitro blue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate to produce a purple precipitate.

Quantitative Real-Time PCR (qRT-PCR). Embryos were infected with the replication-competent retroviral vector (RCAS) RCAS::*Alk5**, RCAS::CA-*β-catenin*, RCAS::*Dkk3*, RCAS::*Bmp4*, and RCAS::CA-*CamKII* constructs and total RNA was extracted from dissected upper beaks of day 11 (st. 33) individuals using an RNeasy kit (Qiagen). Five embryos ($n = 5$) per construct were included in the analysis. RNA was treated with Turbo-DNase (Applied Biosystems) and cDNA was generated using the high-capacity RNA to cDNA kit (Applied Biosystems) and qRT-PCR was performed using the SYBR green protocol (Kapa Biosystems). Forty cycles of amplification were used and data acquisition was carried out with an Eppendorf Mastercycler. We designed specific primers to detect viral infection and all of the genes and skeletogenic markers used in this study (see *SI Materials and Methods* for primer sequences). All primers were designed to detect exclusively the chicken version of each of the genes. Gene expression was assayed in triplicate for each sample and normalized for GAPDH. Data from all qRT-PCR experiments were analyzed using the comparative CT method (1). All levels of expression are reported relative to wild-type (uninfected) embryos. Statistical significant of expression differences was established using a standard two-tailed Student's *t* test.

Quantitative Real-Time PCR Primer Sequences. (*i*) RCAS infection [RCAS-F1 (5′-TCGTTAGCGATGACAATGGA-3′), RCAS-R1 (5′-CACCGAACGTTGTTTGACTG-3′)]; (*ii*) chicken *Bmp4* [*Bmp4*-F1 (5′-ACCATGAAGAGCACCTGGAGAG-3′), *Bmp4*-R1 (5′-TGCTGAGGTTGAAGACGAAGCG-3′)]; (*iii*) chicken *Calmodulin* [*CaM*-F1 (5′-GGCAAGAAAAATGAAAGATA-3′), *CaM*-R1 (5′-GACGAAGTTCTGCAGCACTA-3′)]; (*iv*) chicken *TgfβIIr* [*TgfβIIr*-F1 (5′-AAGGCCTGGGAGAAGAA-TGT-3′), *TgfβIIr*-R1 (5′-GGTTGATGTTGTTGGCACAG-3′)]; (*v*) chicken *β-catenin* [*β-catenin*-F1 (5′-AGGAAGCTGAAATG-GCTCAA-3′), *β-catenin*-R1 (5′-AGATTGCGAATCAACCCAAC-3′)]; (*vi*) chicken Dkk3 [*Dkk3*-F1 (5-GCCCTGTAAAACCCAG-CATA-3′), *Dkk3*-R1 (5′-GAAGTGGCTTTCCTGCACTC-3′)]; (*vii*) *Collagen 2a1* [Bmp4-F1 (5′-AAGGGTGATCGTGGTG-AGAC-3′), Bmp4-R1 (5′-TCGCCTCTGTCTCCTTGTTT-3′)]; (*viii*) *Ostepontin* [*Opn*-F1 (5′-AGCCACCACACACACAGGTA-3′), *Opn*-R1 (5′-TGAAGCCAGGTCATTCTGTG-3′)]; and (*ix*) *GAPDH* [*GAPDH*-F1 (5′-GAGGGTAGTGAAGGCTGCTG-3′), *GAPDH*-R1 (5′-CATCAAAGGTGGAGGAATGG-3′)].

RCAS Constructs. The following plasmids were cloned into RCAS: a constitutively active form of the *TGFβIr* (*Alk-5*) (Addgene; plasmid 14833) (2), a dominant negative form of *TGFβIIr* (Addgene; plasmid 1176) (3), and the entire coding region of the chicken *Dkk3* (gift of Dr. Chrisof Niehrs). RCAS::CA-*β-catenin*, RCAS::*Bmp4*, and RCAS::CA-*CamKII* constructs have been described previously (4–6).

Micro-Computed Tomography (CT) scans. Specimens were scanned using an X-Tek XRA-002 micro-CT imaging system set at 75 kV. Samples were mounted on a rotating table and a series of 3,142 projections of 2,000 × 2,000 pixels covering 360° was recorded. Volume and surface rendering was performed using AMIRA 5.0 (64-bit version; Mercury Computer Systems) for all specimens and the volume of the upper beak was extracted. As species differ in their body and head size, we calculated a multivariate indicator of overall size. To do so we used the $\log_{10}$ transformed wing chord length, tarsus length, head length, head width, and head depth for each specimen as input into a factor analysis, which resulted in a single new factor hereafter referred to as "size." Log_{10} transformed beak volume was then regressed against size and unstandardized residuals were extracted for comparison.

Quantification of Gene Expression Area. Unprocessed raw images were converted into 8-bit format using the program ImageJ (7). The beak profiles were outlined manually to include only the mesenchymal region and exclude the epithelial portion of the tissue. The start of the beak was defined by a line perpendicular to the point where the esophagus region begins (downward curvature). This landmark could be easily identified in all species and stages analyzed, allowing us to have accurate comparisons. For each gene analyzed, a set threshold was chosen and the same value was applied to each of the images for the different species within that particular gene. ImageJ (7) was used to calculate the percentage of the beak where the genes of interest showed expression. Plotted values represent averages (and SD) from three individuals (see Fig. S2 for more details).

Microarray Production and Use. A DNA microarray (21,168 spots) was printed from a nonnormalized poly(A)-primed cDNA library made from RNA isolated from multiple (12 individuals) frontonasal processes of stage-26 and stage-29 embryos of the medium ground finch, *G. fortis* (8). We used Cy5-labeled probes made from individual frontonasal processes of the four derived species of *Geospiza* for direct comparisons against a common Cy3-labeled reference sample made from pooled RNA of several (9 individuals) embryos of more basal *G. difficilis*. In most cases we compared 4 unrelated individuals from each of the derived cactus finch and ground finch species (*G. scandens*, *G. conirostris*, *G. magnirostris*, and *G. fortis*) against the pooled common reference. RNA from each individual finch beak prominence was independently amplified and labeled in triplicate with a control dye swap. We used the two highest-quality sets of microarray data from each triplicate for clustering. Raw .gpr files were generated with GenePix 3.0 (Molecular Devices). Normalization and statistical analysis of the GPR data files were performed in

MatLab (Math Works). Data were normalized with the Lowess algorithm22. Only spots with a signal intensity exceeding the median background +2 SD were considered, which left 7,369 spots. The data were log2 transformed.

Microarray Cluster Analysis. Clustering analysis and visualization were computed in MatLab. Agglomerative hierarchical clustering was performed by using the Euclidean distance measure: the average linkage and Ward heuristics were used to connect the gene clusters. For *k*-means clustering, the *k*-means algorithm partitioned the genes into *k*-discrete clusters on the basis of their expression. The number *k* (50) was preselected. The resultant tree illustrates that duplicates of the amplification/labeling experiments from the same individuals clustered together. Measurements of signal ratios and intensities for different transcripts were clustered to identify genes that were up-regulated or down-regulated in all individuals of a particular species compared with the basal *G. difficilis* reference.

Species-specific clusters were further cross-compared, which reveal transcripts that were consistently up-regulated in frontonasal processes of all individuals of the ground-finch beak morphology and that remained unchanged or were down-regulated in beak primordia of the cactus finches. Conveniently, the cactus finch has an upper beak depth similar in size to that of the medium ground finch, whereas the large cactus finch's beak is more similar in size to that of the large ground finch, although differing strongly in shape. This allowed us to separate transcripts exhibiting the size-specific regulation from those with shape-specific regulation. Gene expression patterns of nine experimental samples representing five *G. scandens* and four *G. conirostris* samples were analyzed by hierarchical clustering with Ward linkage. The samples were divided into two groups on the basis of differences in gene expression. Further details are described in ref 8. Microarray data are available in the ArrayExpress database (www.ebi.ac.uk/arrayexpress) under accession number E-MEXP-702.

1. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25:402–408.
2. Feng XH, Derynck R (1996) Ligand-independent activation of transforming growth factor (TGF) beta signaling pathways by heteromeric cytoplasmic domains of TGF-beta receptors. *J Biol Chem* 271:13123–13129.
3. Wrana JL, et al. (1992) TGF beta signals through a heteromeric protein kinase receptor complex. *Cell* 71:1003–1014.
4. Duprez D, et al. (1996) Overexpression of BMP-2 and BMP-4 alters the size and shape of developing skeletal elements in the chick limb. *Mech Dev* 57:145–157.
5. Kengaku M, et al. (1998) Distinct WNT pathways regulating AER formation and dorsoventral polarity in the chick limb bud. *Science* 280:1274–1277.
6. Taschner MJ, Rafigh M, Lampert F, Schnaiter S, Hartmann C (2008) Ca2+/Calmodulin-dependent kinase II signaling causes skeletal overgrowth and premature chondrocyte maturation. *Dev Biol* 317:132–146.
7. Abramoff MD, Magelhaes PJ, Ram SJ (2004) Image Processing with ImageJ. *Biophotonics International* 11:36–42.
8. Abzhanov A, et al. (2006) The calmodulin pathway and evolution of elongated beak morphology in Darwin's finches. *Nature* 442:563–567.
9. Bowman RI (1961) *Morphological Differentiation and Adaptation in the Galápagos Finches* (Univ of California Press, Berkeley).
10. Grant PR, Abbott I, Schluter D, Curry RL, Abbott LK (1985) Variation in the size and shape of Darwin's finches. *Biol J Linn Soc* 25:1–39.

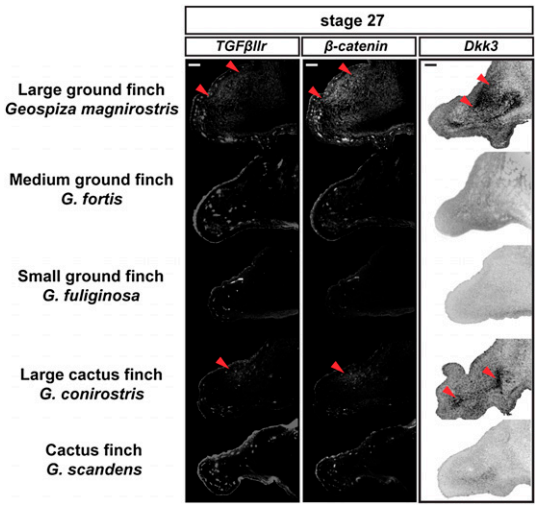

Fig. S1. Variation in the premaxillary bone (pmx) in *Geospiza* correlates with the expression of *TGFβIIr*, *β-catenin*, and *Dkk3*. In the large ground finch, the pmx condensation forms earlier (st. 27) than in the other species, showing a strong correlation with the time and place of expression of *TGFβIIr*, *β-catenin*, and *Dkk3*. Arrow colors indicate species that have comparable body sizes but differ in beak morphology. (Scale bar, 0.1 mm.) Images of skulls are from ref 9, with permission from the author. pmx, premaxillary bone; pnc, prenasal cartilage.

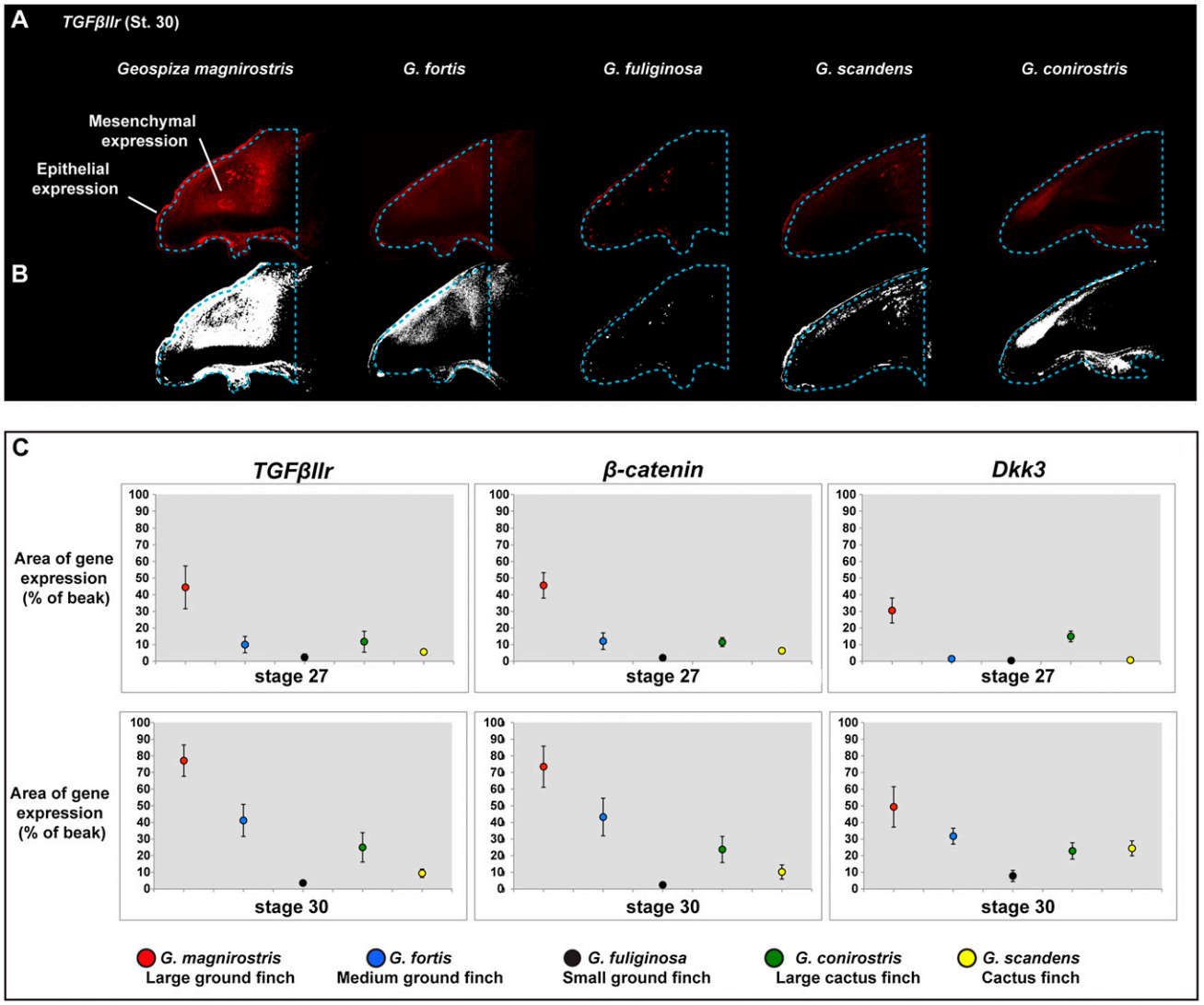

Fig. S2. Quantification of gene expression area. (*A*) Unprocessed raw images of *TGFβIIr* expression in the different species of *Geospiza* analyzed. *TGFβIIr* expression is used here to illustrate the methodology used in this analysis and the same procedure was followed for *β-catenin* and *Dkk3*. Because there is expression in both the beak mesenchyme and the epithelial tissue, beak profiles were outlined manually to include only the mesenchymal region and exclude the epithelial portion of the tissue (blue dashed line). The start of the beak was defined by a line perpendicular to the point where the esophagus region begins (downward curvature). (*B*) Images in 8-bit format were thresholded to the same set value. This value varied between genes but was kept constant for all of the images within a gene. (*C*) For each stage and gene analyzed, the percentage of the beak where the gene was expressed was calculated (error bars represent SD values).

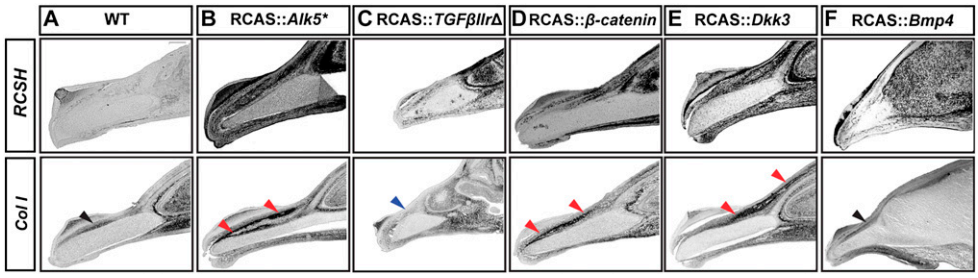

Fig. S3. Functional analysis of *TGFβIIr*, *β-catenin*, and *Dkk3* in the chicken model system. (*A–F*) UV pictures of embryonic day 11 (HH st. 37). (*A*) Wild-type chicken embryos and embryos infected with (*B*) RCAS::*Alk5**, (*C*) RCAS::*TGFβr*Δ, (*D*) RCAS::CA-*β-catenin*, (*E*) RCAS::*Dkk3*, and (*F*) RCAS::*Bmp4* constructs. We used the viral-specific probe RSCH and *Col I* probes to reveal RCAS infection (RSCH) and overall bone (*Col I*). Blue arrows indicate lower expression relative to wild-type specimens, red arrows indicate higher expression, and black arrows indicate no change. (Scale bar, 0.4 mm in sections *A–F*.)

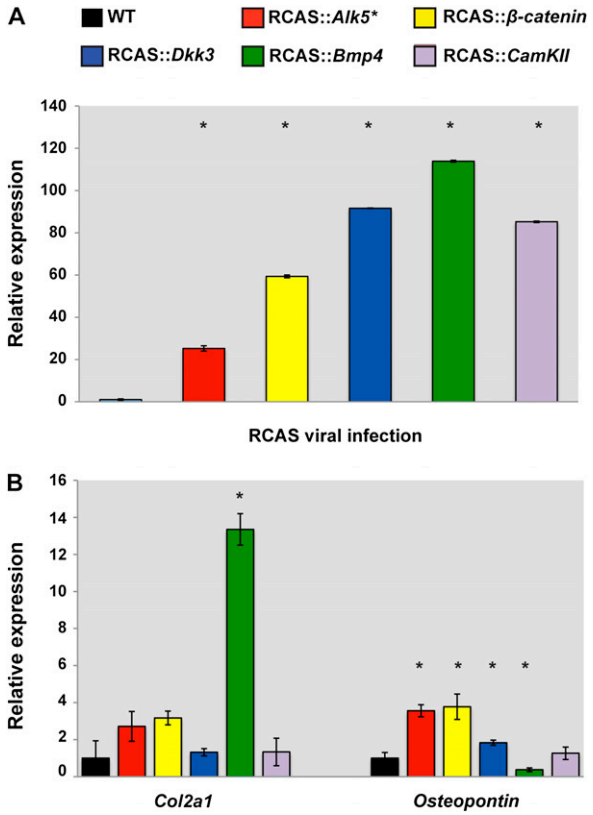

Fig. S4. Quantitative real-time PCR (qRT-PCR) assays of skeletogenic markers. (*A*) qRT-PCR assays showing levels of viral infection in each of the constructs used, as determined by using viral-specific primers. (*B*) Relative expression levels of chondrogenic (*Col2a1*) and osteogenic (*Ostepontin*) molecular markers in beaks of day 11 (HH st. 37) embryos infected with RCAS::*Alk5**, RCAS::CA-*β-catenin*, RCAS::*Dkk3*, RCAS::*Bmp4*, and RCAS::*CamKII* constructs. In embryos infected with RCAS::*Bmp4*, expression of *Col2a1* increases drastically, whereas it remains unchanged in the others. In embryos infected with RCAS::*Alk5**, RCAS::CA-*β-catenin*, and RCAS::*Dkk3*, *Osteopontin* levels increase relative to uninfected controls, whereas they significantly decrease in chicks infected with RCAS::*Bmp4*. Expression levels are shown relative to wild-type uninfected controls (asterisks denote significance at $P < 0.05$, *t* test, $n = 5$; error bars represent SD values).

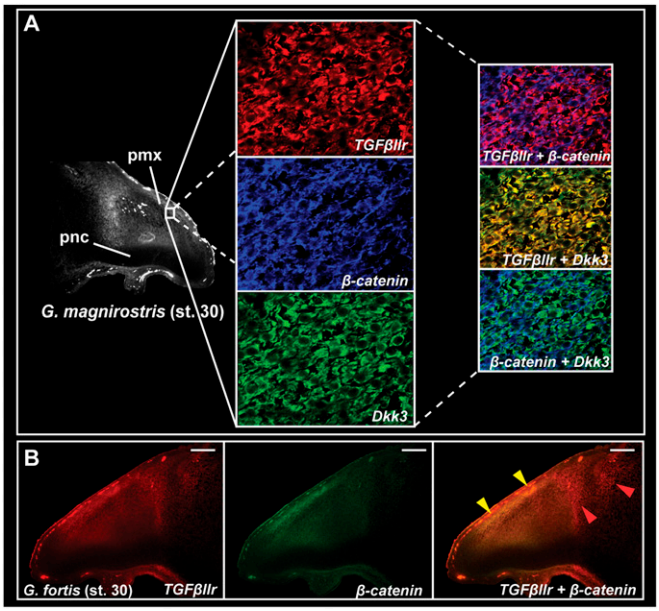

Fig. S5. Confocal microscopy image of developing Darwin's finch beaks stained for the genes expressed in the premaxillary bone. (*A*) At st. 30, *TGFβIIr*, *β-catenin*, and *Dkk3* can be coexpressed in the same cells in the developing mesenchyme. (*B*) However, the three molecules are not necessarily expressed in the same cells. Sagittal sections of a stage-30 *G. fortis* stained for *TGFβIIr* and *β-catenin* show that some mesenchymal cells coexpress both molecules (yellow arrows in merged image, whereas other cells only express *TGFβIIr* (red arrows). (Scale bar, 0.2 mm in *B*.) pmx, premaxillary bone; pnc, prenasal cartilage.

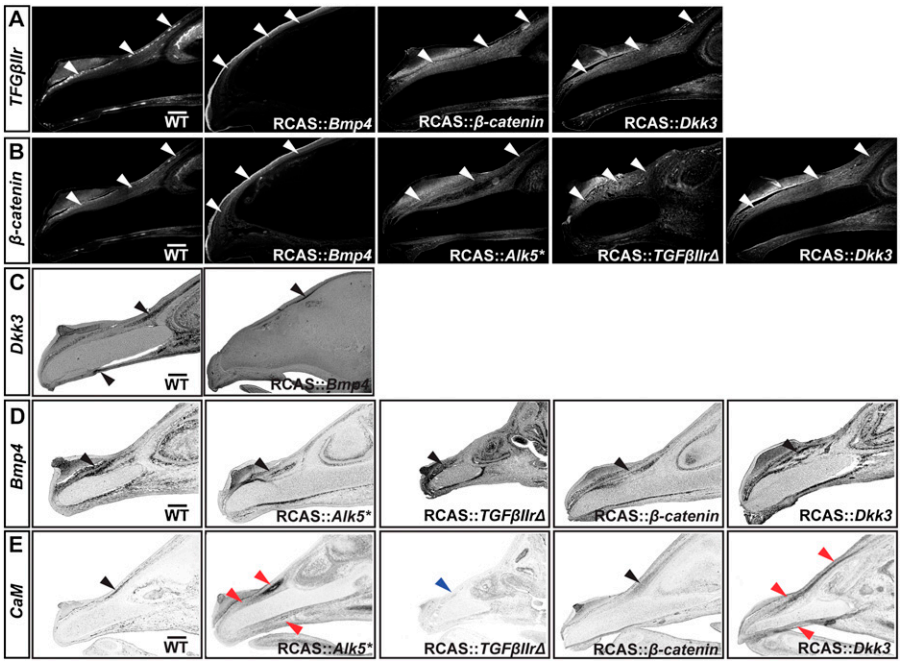

Fig. S6. *Bmp4*, *CaM*, *TGFβIIr*, *β-catenin*, and *Dkk3* expression patterns in beaks of day-11 (HH st. 37) embryos infected with RCAS::*Alk5**, RCAS::CA-*β-catenin*, RCAS::*Dkk3*, RCAS::*Bmp4*, and RCAS::*CamKII* constructs. (*A*) Expression of chicken *TGFβIIr* did not change in any of the infected embryos relative to uninfected controls. (*B*) Similarly, chicken *β-catenin* remained consistent across embryos infected with different constructs. (*C*) Expression of *Dkk3* remained unchanged in RCAS::*Bmp4* infected embryos. (*D*) Similarly, expression of *Bmp4* remained unchanged across treatments, whereas (*E*) expression of *CaM* significantly increased when TGFβ signaling and *Dkk3* were up-regulated but remained unchanged when *β-catenin* levels were up-regulated. Conversely, down-regulation of TGFβ signaling led to a decrease in *CaM.*

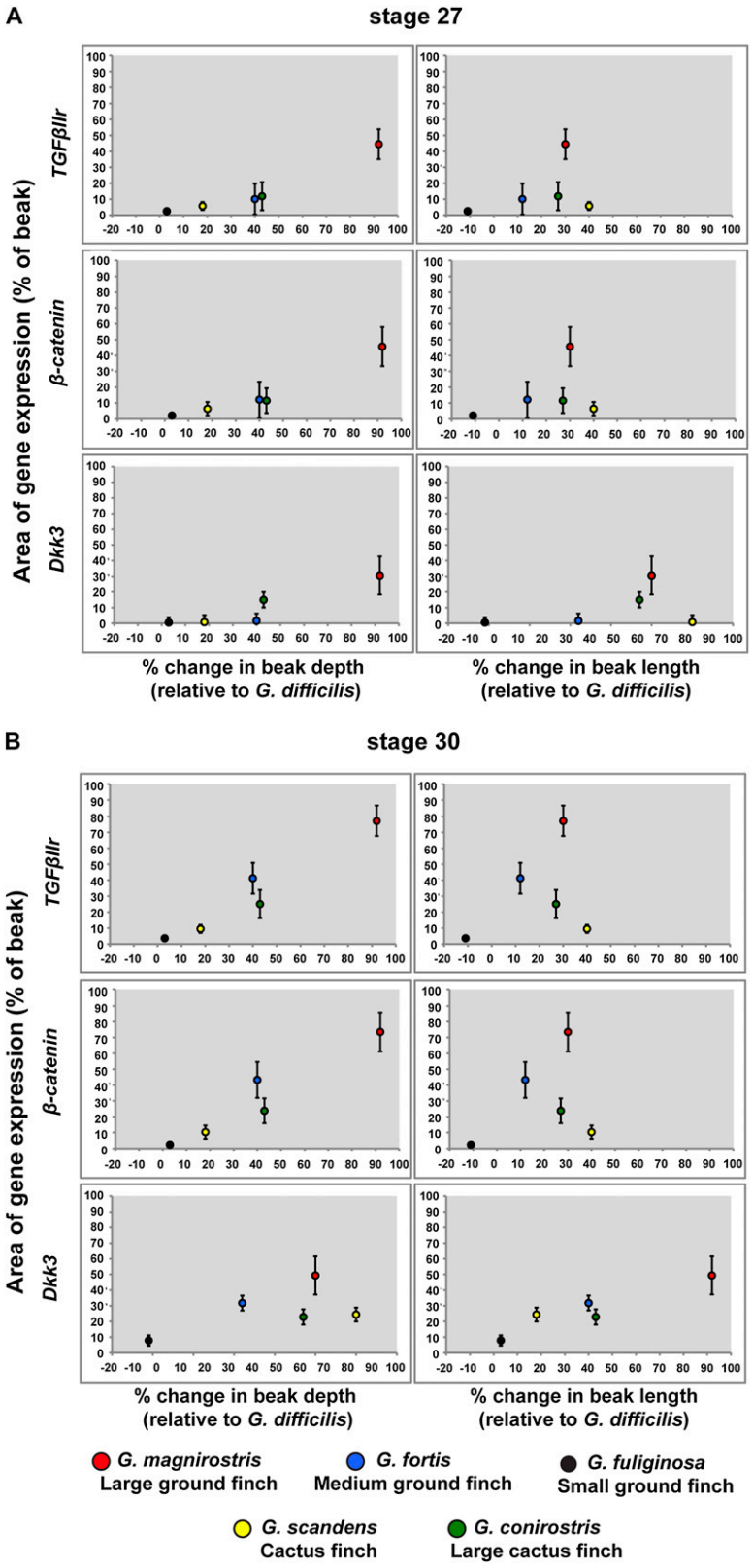

Fig. S7. Quantification of gene expression area in st.-27 (*A*) and st.-30 (*B*) Darwin's finch embryos and its correlation with adult beak morphology. Images where processed as described in Fig. S3 and in *SI Materials and Methods*. Measurements in *C* were taken from ref. 10, corrected for wing length, and correspond to averages from males that were collected in the islands where we obtained our samples (error bars represent SD values).

Table S1. Premaxillary bone volume of members of the genus *Geospiza* as determined with (micro) computed tomography (CT) scans

Species	Volume (mm^3)	Size corrected (residual) volume
Geospiza difficilis	18.8	−0.378
Geospiza fuliginosa	34.2	0.253
Geospiza fortis	112.0	0.203
Geospiza magnirostris	312.3	0.234
Geospiza conirostris	78.66	−0.092
Geospiza scandens	49.3	−0.219

As species differ in their body and head size, we calculated a multivariate indicator of overall size. To do so we used the $\log_{10}$ transformed wing chord length, tarsus length, head length, head width, and head depth for each specimen as input into a factor analysis, which resulted in a single new factor hereafter referred to as "size." Log_{10}-transformed beak volume was then regressed against size and unstandardized residuals were extracted for comparison.

Table S2. Final cluster of ground finch-specific genes

No.	Clone no.	Clone identity from ground finch-specific cluster	Ground/cactus fold difference	Expression signal intensity
1	0019H10	Unknown	17.218	5,062
2	0041H10	XDRP1(binds to cyclinA)	18.007	3,031
3	0017D7	Unknown	10.456	2,698
4	0071G4	Proteasome subunit, alpha type 7	18.735	120
5	0063C4	Proteasome subunit, alpha type 5	10.517	10,494
6	0055A4	Unknown	20.533	725
7	0062B1	20S proteasome alpha 4 subunit	19.063	3,159
8	0074 e10	Ribosomal protein L23 (L17)	16.043	669
9	0074G1	Zinc finger protein 335	15.952	115
10	0074G4	Unknown	12.325	5,136
11	0014 E7	Proteasome subunit, alpha type 7	11.035	4,868
12	0096F11	Similar to ribosomal protein L23	16.757	479
13	0022A2	Unknown	12.716	8,490
14	0051C9	Unknown	8.801	11,132
15	0044F3	Similar to ribosomal protein L23	13.322	170
16	0044F6	Unknown	14.222	2,257
17	0044F9	Unknown	12	6,513
18	**0181F7**	**TGF β receptor type II**	**10.333**	**7,698**
19	0181F10	Unknown	9.25	3,434
20	0116B10	Unknown	11.847	1,013
21	0212 E10	Unknown	10.75	3,710
23	0137 E5	Unknown	14.599	741
22	0154C10	Ribosomal protein L23	9.417	1,020
24	**0164F2**	**Dickkopf-3 (DKK-3)**	**6.867**	**17,167**
25	0212F11	Unknown	6.482	6,915
26	0178B2	Ribosomal protein L23	8.867	1,705
27	0118F5	UNKNOWN	6.914	43,115
28	0178B5	AMPK alpha-1 chain	6.962	10,243
32	0118F8	Unknown	3.003	30,638
30	0178B8	Unknown	14.333	6,138
31	0130D8	Unknown	4.818	414
32	0192 E5	Ribosomal protein L23	5.246	192
33	0192 E8	Unknown	3.445	2,408
34	0120A8	GTP-binding protein RAB1A	3.336	17,747
35	0130G8	Unknown	3.147	3,168
36	0127B3	Ribosomal protein L23	5.02	1,220
37	0190 E6	Unknown	4.935	31
38	0190 E9	Ribosomal protein L18a	8.821	15,683
39	0097F1	Unknown	14.273	8,729
40	0107B11	Ribosomal protein S11	10.5	4,165
41	0029B8	TRAM-1	27.5	4,238

Table S2. Cont.

No.	Clone no.	Clone identity from ground finch-specific cluster	Ground/cactus fold difference	Expression signal intensity
42	0025D5	Unknown	8.729	11,636
43	0064B5	Unknown	35.667	14,487
44	0006D2	Unknown	7.124	14,169
45	0147B1	Unknown	8.212	5,862
46	0195H1	Nucleoside diphosphate kinase	8.756	3,461
47	0157H4	40S ribosomal protein S3	10.771	2,911
48	0149D4	Nucleoside diphosphate kinase	4.407	16,107
49	0187 E1	Histone H3, embryonic	9.008	5,257
50	0187 E4	Unknown	12.444	862
51	0187 E7	Alpha-cardiac actin	9.177	309
52	0143 E7	Unknown	4.694	750
53	0125G7	Nucleoside diphosphate kinase	3.533	15,443
54	0185C10	40S ribosomal protein S21	3.992	19,207
55	0185 E1	Unknown	3.233	6,434
56	0129A4	XP_237708	5.051	4,125
57	0185 E4	Unknown	8.797	3,637
58	0129A7	Translation initiation factor 3, subunit 3	9.234	1,056
59	0181C4	Unknown	7	8,291
60	0129A10	Unknown	11.746	340
61	0193 E7	Unknown	10.39	2,866
62	0152B10	TRAM-1	12.627	221
63	0204D10	Nucleoside diphosphate kinase	12.15	69
64	0152F4	Ribosomal protein S11	7.622	5,223
65	**0132F10**	**Beta-catenin (β-catenin)**	**4.281**	**7,478**
66	0194D7	Unknown	18.6	88
67	0212 E4	ETS-family transcription factor EHF	7.353	32,394
68	0144 E7	Unknown	5.8	24,120
69	0162 E1	Peptide elongation factor 1-beta	14.254	11,288
70	0162 E10	Ribosomal protein S11	11.143	1,472
71	0210G1	Unknown	19.5	8
72	0159H8	Diphosphate kinase	6.03	17,609
73	0179B2	Unknown	9.464	2,699
74	0179B8	Unknown	8.985	16,087
75	0167F5	Ribosomal protein L35a	16.25	3,353
76	**0211F8**	***Coll Ia2* precursor (osteoblast marker)**	7.476	6,305
77	0165B2	Unknown	6.32	13,504
78	0113H8	Peptide elongation factor 1-beta	6.107	12,298
79	0189D5	Unknown	18.212	888
80	0189D8	Unknown	9.756	1,332
81	0189D11	Unknown	10.771	8,132
82	0153D2	40S ribosomal protein S21	10.407	2,503
83	0131C2	Ribosomal protein S11	16.008	294
84	0173A5	Ribosomal protein S21	4.444	5,063
85	0173A8	Unknown	9.177	867
86	0125 E8	Unknown	4.694	8,341
87	0125 E11	Unknown	5.533	1,243
88	0125G2	Unknown	3.992	12,755
89	0208B11	GALECTIN-1	3.233	8,937
90	0110H8	Unknown	3.051	8,754
91	0194H11	Unknown	4.639	8,246
92	0130D11	Unknown	5.961	5,217
93	0114F5	Unknown	7.908	956
94	0200A11	Peptide elongation factor 1-β	5.904	167
95	0188G5	Unknown	6.343	2,300
96	0188G8	Unknown	6.294	20,844
97	0097F1	Unknown	6.086	7,969
98	0107B11	Ribosomal protein S11	5.269	4,445
99	0029B8	TRAM-1	4.25	7,394
100	0025D5	Unknown	3.492	8,638

Both the clone number as found in the library as well as the identity as revealed by sequencing and BLAST search analysis are shown. The table shows a cluster, which showed candidates with *both* high median-fold difference in expression between ground and cactus finches and overall expression (signal) level. The three genes examined here, *TGFβIIr* (no. 18), *Dkk3* (no. 24), and *β-catenin* (no. 65), are highlighted in bold. Also, note that the array revealed an important early bone marker, alpha-2 type I collagen precursor (*Coll Ia*) (no. 76), also highlighted in bold.