Independent regulation of vertebral number and vertebral identity by microRNA-196 paralogs

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

| As Published | http://dx.doi.org/10.1073/pnas.1512655112 |
| Publisher | National Academy of Sciences (U.S.) |
| Version | Final published version |
| Accessed | Wed Mar 16 09:10:30 EDT 2016 |
| Citable Link | http://hdl.handle.net/1721.1/101423 |
| Terms of Use | Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use. |
| Detailed Terms | |
Independent regulation of vertebral number and vertebral identity by microRNA-196 paralogs

Siew Fen Lisa Wong\textsuperscript{a,1}, Vikram Agarwala\textsuperscript{b,c,d,e,1}, Jennifer H. Mansfield\textsuperscript{a,1}, Nicolas Denans\textsuperscript{b}, Matthew G. Schwartz\textsuperscript{f}, Haydn M. Prosser\textsuperscript{1}, Olivier Pourquié\textsuperscript{e,1}, David P. Bartel\textsuperscript{b,c,d}, Clifford J. Tabin\textsuperscript{1,2}, and Edwina McGlinn\textsuperscript{a,1,2}

\textsuperscript{a}EMBL Australia, Australian Regenerative Medicine Institute, Monash University, Clayton, VIC 3800, Australia; \textsuperscript{b}Howard Hughes Medical Institute, Cambridge, MA 02142; \textsuperscript{c}Whitehead Institute for Biomedical Research, Cambridge, MA 02142; \textsuperscript{d}Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139; \textsuperscript{e}Computational and Systems Biology Program, Massachusetts Institute of Technology, Cambridge, MA 02139; \textsuperscript{f}Department of Genetics, Harvard Medical School, Boston, MA 02115; \textsuperscript{g}Department of Biological Sciences, Barnard College, New York, NY 10027; \textsuperscript{h}Department of Developmental Biology and Genetics, Stanford School of Medicine, Stanford, CA 94305, The Welcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, United Kingdom; and \textsuperscript{i}Department of Pathology, Brigham and Women's Hospital, Boston, MA 02115

Contributed by Clifford J. Tabin, July 16, 2015 (sent for review March 24, 2015; reviewed by Jacqueline Deschamps and Joshua T. Mendell)

The Hox genes play a central role in patterning the embryonic anterior-to-posterior axis. An important function of Hox activity in vertebrates is the specification of different vertebral morphologies, with an additional role in axis elongation emerging. The miR-196 family of microRNAs (miRNAs) are predicted to extensively target Hox expression dynamics and influences mammalian development remains to be elucidated. Here we used an extensive allelic series of mouse mutants over more than 20 years to analyze miR-196 paralogs, 196a1, 196a2, and 196b act redundantly to constrain total vertebral number. Loss of miR-196 leads to a collective up-regulation of numerous trunk Hox target genes with a concomitant delay in activation of caudal Hox genes, which are proposed to signal the end of axis extension. Additionally, we identified altered molecular signatures associated with the Wnt, Fgf, and Notch/segmentation pathways and demonstrate that miR-196 has the potential to regulate Wnt activity by multiple mechanisms. By feeding into, and thereby integrating, multiple genetic networks controlling vertebral number and identity, miR-196 is a critical player defining axial formulae.

\textbf{Hox genes | microRNA | miR-196 | axial patterning | vertebral specification}

A defining feature of the vertebrate animals is the presence of a segmented vertebral column. Species are uniquely characterized by the total number of vertebrae that form and by the regionalization of these vertebrae along the anterior-to-posterior axis into groups with distinct morphologies (e.g., cervical, thoracic, lumbar, and sacral). The genetic determinants of vertebral number and vertebral identity have largely been considered as separate; thus, how, or even whether, these processes are molecularly integrated remains to be clearly elucidated.

Vertebral precursors, known as somites, arise by continued expansion and segmentation of a region of the caudal embryo, the presomitic mesoderm (PSM) (1). Expansion of the PSM requires a self-renewing axial progenitor population that initially resides in the node-streak border of the epiblast and subsequently repositions to the tailbud (2–6). These progenitors provide a source of cells that, following ingress through the primitive streak, populate the PSM and other derivatives to drive posterior elongation. Key players in this process include genes involved in Wnt and Fgf signaling, in addition to the Cdx transcription factors, as evidenced by severe axis truncations when each are mutated (7).

Balancing the expansion of this cell population, cells of the anterior PSM bud off to form somites with a rhythmic periodicity inherent to each species. The eventual exhaustion of progenitor self-renewal capacity is thought to halt axis elongation, the timing of which is a critical factor in establishing species-specific vertebral number (8).

Within vertebral precursors, specific combinations of Hox transcription factors impart positional information that governs vertebral identity (9). In mammals, the 39 Hox genes are clustered at four separate genomic loci (HoxA, HoxB, HoxC, and HoxD), with each gene classified into 1 of 13 paralogous groups dependent on sequence similarities and relative positions within the respective clusters (Fig. S4). These genes are expressed in partially overlapping domains during embryonic development, with a spatiotemporal collinearity that reflects genomic ordering (10, 11). Exhaustive analysis of Hox mouse mutants over more than 20 years has revealed individual and cumulative Hox function in conferring specific positional identities to the forming vertebral column (9). For instance, the central/trunk Hox proteins (paralogs 5–8) primarily pattern thoracic vertebra, whereas Hox 11 paralogs pattern sacral and caudal vertebra (12) and position the sacrum (13, 14).

In addition to transcripts encoding for the Hox proteins, transcription within the genomic Hox clusters produces noncoding regulatory RNAs, including several microRNAs (miRNAs) (Fig. S4) (15). In mice, these include the miR-10 family, which is found to self-renewal capacity is thought to halt axis elongation, the timing of which is a critical factor in establishing species-specific vertebral number (8).

Within vertebral precursors, specific combinations of Hox transcription factors impart positional information that governs vertebral identity (9). In mammals, the 39 Hox genes are clustered at four separate genomic loci (HoxA, HoxB, HoxC, and HoxD), with each gene classified into 1 of 13 paralogous groups dependent on sequence similarities and relative positions within the respective clusters (Fig. S4). These genes are expressed in partially overlapping domains during embryonic development, with a spatiotemporal collinearity that reflects genomic ordering (10, 11). Exhaustive analysis of Hox mouse mutants over more than 20 years has revealed individual and cumulative Hox function in conferring specific positional identities to the forming vertebral column (9). For instance, the central/trunk Hox proteins (paralogs 5–8) primarily pattern thoracic vertebra, whereas Hox 11 paralogs pattern sacral and caudal vertebra (12) and position the sacrum (13, 14).

In addition to transcripts encoding for the Hox proteins, transcription within the genomic Hox clusters produces noncoding regulatory RNAs, including several microRNAs (miRNAs) (Fig. S4) (15). In mice, these include the miR-10 family, which is found to act redundantly to constrain total vertebral number. Loss of miR-196 leads to a collective up-regulation of numerous trunk Hox target genes with a concomitant delay in activation of caudal Hox genes, which are proposed to signal the end of axis extension. Additionally, we identified altered molecular signatures associated with the Wnt, Fgf, and Notch/segmentation pathways and demonstrate that miR-196 has the potential to regulate Wnt activity by multiple mechanisms. By feeding into, and thereby integrating, multiple genetic networks controlling vertebral number and identity, miR-196 is a critical player defining axial formulae.

Significance

The vertebral column provides essential structural and protective functions. The total number of vertebral elements and their specific morphologies are remarkably reproducible within a given species, yet can be tailored to the requirements of separate vertebrate species. Major genetic determinants driving formation of the vertebral column are known, but how they are regulated to achieve a highly reproducible structure remains to be fully elucidated. In this report, we show that the miR-196 family of microRNAs are essential in defining correct vertebral number and vertebral identity in mouse. We reveal the molecular landscape controlled, either directly or indirectly, by miR-196 activity, to demonstrate that miR-196 impacts many key developmental signalling pathways and reinforces a timely trunk-to-tail Hox code transition.


Reviewers: J.D., Hubrecht Institute; and J.T.M., UT Southwestern Medical Center. The authors declare no conflict of interest.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE53018).

1S.F.L.W. and V.A. contributed equally to this work.

2To whom correspondence may be addressed. Email: tabin@genetics.med.harvard.edu or edwina.mcglinn@embalgaustalia.org.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1512655112/-/DCSupplemental.
appears to be an “insertion” of a rib-bearing precaudal element (24). How this latter phenotype arises developmentally is not known and is difficult to reconcile with derepression of trunk Hox target genes alone (31, 32). These knockdown approaches could not shed light on individual paralog contributions for this highly related miRNA family, and importantly, the molecular networks downstream of miR-196, which have the potential to drive phenotypic alterations, remain uncharacterized.

Here, we generated individual KO alleles for each of the three miR-196 family members in mice. These tools have allowed us to build an entire allelic deletion series to reveal the individual and additive roles of miR-196 paralogs in patterning vertebral identity at many axial levels and in controlling the total number of vertebrae. We characterized the detailed molecular landscape controlled by miR-196 activity in the early embryo to show that miR-196 regulates, and therefore has the ability to integrate, multiple key signaling pathways to drive developmental processes.

Results
Differential Transcription of miR-196a1 and miR-196a2 in the Developing Embryo. To reveal the individual expression patterns, and therefore potential for functional redundancy, of identical miRNAs 196a1 and 196a2, we generated eGFP knock-in alleles termed 196a1/GFP and 196a2/GFP (Fig. S1). Expression of reporter miRNA reflects sites of active transcription, although it does not reveal additional posttranscriptional regulation that endogenous miRNAs may undergo. Whole mount in situ hybridization analysis of reporter miRNA indicated that both miRNAs were expressed specifically in the posterior embryonic derivatives of all three germ layers and revealed striking differences in their spatiotemporal kinetics that have not previously been delineated (Fig. 1 B and C). miR-196a1 is expressed throughout the posterior growth zone at embryonic day (E)8.5 (Fig. 1B). Strong expression is maintained in the PSM until the end of axis elongation, with a discrete band of low expression in the anterior PSM from E10.5 (Fig. 1B, Inset). The anterior boundary of somitic and neural expression extends to approximately somite 13/14 [prevertebra (pv) 9, thoracic (T) 2] at E9.5 with a caudal shift in somitic tissue and a rostral shift in neural tissue as development proceeds. This expression profile indicates that 196a1 exhibits a classic collinear profile relative to the adjacent Hox gene, Hoxb9 (Fig. 1B, Inset). miR-196a2 expression is temporally delayed relative to 196a1 (Fig. 1C), with faint expression ventral to the PSM at E8.5–9.0. Strong expression is then observed throughout the PSM and neural plate at E9.5. A stable anterior somitic limit at approximately somite 21/22 (pv17, T10) and neural limit 2 somites rostral to this is established soon after, consistent with its positioning between Hox9 and Hox10 (34). This analysis revealed both unique and overlapping expression patterns of 196a1 and 196a2, suggesting these identical miRNAs might have both unique functions where individually expressed and either redundant or additive functions at sites of coexpression.

Genetic Deletion of miR-196 Leads to Altered Vertebral Identity. The collective function of miR-196 family members has yet to be assessed in mammals. Moreover, the dissection of paralog contributions to overall miR-196 activity has not been achieved in any system. To address this, we generated straight KO alleles at each of the three murine miR-196 paralogs (24, 33). miR-196a2 expression is temporally delayed relative to 196a1 (Fig. 1C), with faint expression ventral to the PSM at E8.5–9.0. Strong expression is then observed throughout the PSM and neural plate at E9.5. A stable anterior somitic limit at approximately somite 21/22 (pv17, T10) and neural limit 2 somites rostral to this is established soon after, consistent with its positioning between Hox9 and Hox10 (34). This analysis revealed both unique and overlapping expression patterns of 196a1 and 196a2, suggesting these identical miRNAs might have both unique functions where individually expressed and either redundant or additive functions at sites of coexpression.

Removal of individual miR-196 paralogs alone revealed partially penetrant homeotic patterning defects (Fig. 2A and Table S1). In 196a2 or 196b single-mutant embryos, the presence of an ectopic rudimentary rib nubbin on the first lumbar vertebra indicated an anterior homeotic transformation of this element (Fig. 2A).
Additionally, in approximately one quarter of cases, we observed anterior homeotic transformations encompassing all subsequent lumbar and sacral elements, resulting in a posterior displacement of the sacrum (schematized in Fig. 2D). Although this latter phenotype could be interpreted as an insertion of a thoracic element, the repositioned last lumbar vertebrae (L6* in Fig. 2D) was often asymmetric, with both lumbar and sacral characteristics (Table S1), which supports the interpretation of serial identity changes, beginning at L1 and encompassing all subsequent elements. We did not observe a similar L1-to-T anterior homeotic transformation in 196aI single-mutant embryos, which for the most part exhibited no overt vertebral alterations (Fig. 2A). However, at very low penetrance (Table S1), 196aI single-mutant embryos displayed an anterior displacement of the sacrum, with or without a reduction

![Fig. 2. miR-196 paralogs function in establishing vertebral identity and number in mouse. (A) Identification of vertebral patterning defects in individual and compound miR-196 loss-of-function E18.5 embryos. Genotypes indicated. The positions of the 13th thoracic element (T13) and first sacral element (S1) are labeled. (Inset) Thoraco-lumbar junction. (B) Individual vertebra analysis to demonstrate identity alterations at the thoraco-lumbar and lumbo-sacral junctions. Genotypes indicated. The position of a rib-like nubbin on lumbar elements is marked with arrow. The position of sacral process is marked with an asterisk. (C) Rib fusion defects observed following loss of miR-196 alleles, genotypes indicated. Fusion of the eighth rib to the sternum was unilateral or bilateral as indicated with arrows. (D) Summary of patterning defects identified across the miR-196 allelic series. An asterisk indicates a homeotic transformation of that vertebral element. (E) Quantification of vertebral number in single and compound miR-196 loss-of-function E18.5 embryos identifies a role for miR-196 in controlling axis length in mouse. Statistical comparison of vertebral number relative to WT was performed using a permutation test, with P values corrected for multiple hypothesis testing using the Bonferroni method: *P < 0.05, ***P < 0.001, and ****P < 0.0001.](http://www.pnas.org/cgi/doi/10.1073/pnas.1512655112)
in rib length of the last thoracic element (T13), suggesting these paralogs may have an opposing role at this axial level.

We hypothesized that the penetrance and severity of the phenotypes observed after mutating single miR-196 paralogs could be enhanced by combining these mutations. Indeed, 196a2−/−;196b−/− double-mutant skeletons exhibited a fully penetrant phenotype, with two pairs of supernumerary ribs and anterior homeotic transformation of all subsequent elements (Fig. 2 A, B, and D). Relative to this double mutant phenotype, triple KO embryos, 196a1−/−;196a2−/−;196b−/−, displayed no additional patterning defects (Fig. 2 A and D, complete dataset summarized in Fig. S3).

We also hypothesized that combining these mutations might reveal additional defects not observed in single mutants. Indeed, all double-mutant skeletons, or skeletons with a triple KO combination of four or more alleles removed, exhibited a partially penetrant increase in the number of ribs attached to the sternum (Table S1 and Fig. 2C), indicating a transformation of the eighth thoracic element to a more anterior identity. Together, our analysis showed that (i) 196a2 and 196b have single and additive effects in patterning the thoraco-lumbar transition and in positioning the sacrum, with a possible opposing role or 196a1 at this axial level, and (ii) 196a1, 196a2, and 196b act redundantly to pattern the midthoracic region, with phenotypic alterations observed only when two or more paralogs are removed. As such, our work has provided the first genetic proof, to our knowledge, for miR-196 as a homeotic family of genes and revealed identity changes at multiple axial levels.

**Genetic Deletion of miR-196 Leads to an Increase in Vertebral Number.** Homeotic transformations do not alter the number of vertebrae, simply their identity. It was therefore surprising that, in zebrafish, miR-196 has been shown to constrain total vertebral number (24). We assessed whether this was an evolutionarily conserved function of miR-196 and found that the three murine miR-196 paralogs constrain total vertebral number in a redundant fashion. Wild-type (WT) C57BL6/J mice exhibit small variations in the total number of vertebrae (Fig. 2E). Compared with the WT mean, we observed a statistically significant increase of approximately one vertebral element in various allelic combinations, including 196a1−/−;196a2−/−, 196a2−/−;196b−/−, and triple KO combinations with four or more alleles deleted (Fig. 2E). Depending on the exact allelic combination, this additional element was patterned as a thoracic (e.g., in 196a2−/−;196b−/− mice) or a post-sacral (e.g., in 196a1−/−;196a2−/− mice) element. Together, these results indicate that miR-196-mediated control of vertebral number and patterning of segment identity are separable processes. All three miR-196 paralogs contribute additively to establishing vertebral number within mouse. Layered on top of this, individual miR-196 paralogs have a differential impact on positional identity and ultimate axial formulae, likely as a result of their differential spatiotemporal kinetics (Fig. 1 B–K) relative to target mRNAs.

**Transcriptome Alterations Are Detected Following Allelic Removal of miR-196 Activity.** To elucidate the molecular mechanism and targets downstream of miR-196, we examined the response of mRNAs to the loss of murine miR-196 alleles in E9.5 embryos. To focus these molecular analyses on the relevant cells, i.e., those cells that normally express miR-196, we used only embryos with at least one eGFP knock-in allele and performed RNA-seq on RNA isolated from cells that were GFP positive (Fig. 3A). With mRNA profiled across 10 genotypes (Table S2), we then compared mRNA

![Fig. 3. Whole transcriptome analysis of miR-196 mutant cells reveals a dysregulation of miRNA targets and skeletal genes. (A) Overview of the experimental and computational strategy used to identify global transcriptome alterations following loss of miR-196 function. (B) Mean fold changes of genes associated with predicted targets of miR-196, partitioned into four context+ intervals according to predicted miRNA targeting efficacy (0 < context+ < −0.2, n = 2,112; −0.2 ≤ context+ < −0.3, n = 145; −0.3 ≤ context+ ≤ −0.4, n = 50; context+ ≤ −0.4, n = 37), across seven genotype comparisons. Statistical comparison of observed up-regulation of genes relative to genes with no miRNA target site, as evaluated by a one-sided Kolmogorov–Smirnov (K-S) test; *P < 0.05, **P < 0.001. (C) Top 10 significant categories related to gene development and function associated with differentially expressed genes. (D) Top 15 categories related to skeletal and muscular development activated in the 196a2−/−;196b−/− vs. 196a2−/− comparison, with corresponding activation z-scores and P values. An activation z-score is a measurement of the consistency between the observed pattern of up- and down-regulation of genes in a category and the predicted activation or inhibition pattern in networks stored in the Ingenuity Knowledgebase relative to a random pattern (B3). P values in C and D are Benjamini–Hochberg corrected P values, with dashed black lines indicating a significance threshold of 0.01.**

Wong et al.  
PNAS  |  Published online August 17, 2015 | E4887
changes as increasing numbers and differing combinations of alleles were deleted (Dataset S1). We first examined the effect of allelic miR-196 deletion on predicted miR-196 target genes. Using the total context score from TargetScan 6.2, which considers the number and type of miRNA binding sites, as well as additional features to predict the genes most effectively targeted by each miRNA (35), we observed that the top predicted targets of miR-196 exhibited significant derepression on the loss of additional miR-196 alleles (Fig. 3B and Fig. S4). The derepression of these predicted targets increased with the number of additional alleles deleted (Fig. 3B), revealing miR-196 dosage sensitivity. The direct interaction between miR-196 and its target transcripts could occur in any of the three germ layer derivatives in which miR-196 was expressed, and indeed, an unbiased analysis of all differentially expressed genes revealed statistically altered molecular signatures (Fig. 4). Of particular interest, we observe statistical enrichment in genes controlling skeletal morphology (Fig. 3C and D and Fig. S5), indicating the presence of a molecular signature consistent with the vertebral abnormalities observed at the phenotypic level.

**Hox Cluster Expression Dynamics Are Altered in miR-196 Mutant Embryos.** It was not known exactly how many of the 10 predicted murine miR-196 Hox target genes are in fact bona fide targets in an in vivo developmental context, nor was it known the relative level of regulation that these predicted targets undergo. When specifically interrogating our transcriptome datasets to assess effects on Hox gene expression, a significant and dose-dependent up-regulation of predicted miR-196 Hox targets was observed (Fig. 4A), which paralleled the dose-dependent patterning defects (Fig. 24). Comparison of 196a2+/−;196b−/− vs. 196a2+/− profiles identified 7 of 10 predicted miR-196 Hox targets as significantly derepressed in double-mutant cells at this developmental stage. Those predicted Hox targets exhibiting no significant derepression in our analysis included Hoxb1, Hoxa4, and Hox5. The most highly derepressed Hox targets were Hoxc8 and Hoxa7, both of which harbor multiple predicted miR-196 binding sites in their 3′ UTRs, and Hoxb8, which exhibits unusually extensive complementarity to miR-196 (17, 23). Further, the measurement of differential expression (Fig. 4A) was almost certain to be an underestimate, because our strategy used eGFP-positive control samples in which at least one miR-196 allele had been removed. Whole mount in situ hybridization (WISH) further revealed that the derepression of Hoxb8 and Hoxc8 target transcripts in 196a2+/−;196b−/− embryos relative to WT manifested as a posterior expansion of Hox8 (B; n = 3/3) and Hoxc8 (C; n = 3/3). The PSM is indicated with a red line/arrowhead and neural tube with a white arrowhead.

![Fig. 4. Loss of miR-196 function alters global Hox signatures. (A) Extensive Hox gene dysregulation is identified following loss of miR-196. Quantitative expression analysis of all 39 Hox genes in cells isolated from E9.5 mutant embryos; genotype comparisons are color coded. Hox genes with one or more predicted miR-196 target binding sites are indicated in red. Filled circles at the tips of fold changes represent a statistically significant change at q < 0.05. (B and C) WISH analysis of miR-196a2−/−;miR-196b−/−. E9.5 embryos relative to WT identifies a caudal expansion of Hox8 (B; n = 3/3) and Hoxc8 (C; n = 3/3). The PSM is indicated with a red line/arrowhead and neural tube with a white arrowhead.](www.pnas.org/cgi/doi/10.1073/pnas.1512655112)
absence of predicted miR-196 sites within these posterior Hox mRNAs, together with the direction of the regulation (down instead of up with diminished miRNA), indicated that this regulation (i.e., have a context+ score ≤ −0.2). Genes with one or more conserved miR-196 target binding sites are indicated in green. (B) In vitro luciferase analysis confirms sequence-specific regulation of three experimentally supported target genes of miR-196. Renilla luciferase intensity values have been normalized to their respective Firefly values (RLU). Controls (WT 3′ UTR construct without miR-196b) were set to 1. MUT, mutated 3′ UTR construct destroying miR-196 binding site. Error bars represent SD. P values, Student t test: *P < 0.05, ***P < 0.0005, and ****P < 0.0001.

Identification of Additional Direct Targets of mir-196. The statistical enrichment of Hox genes among all miR-196 predicted targets (18) prioritized these mRNAs for immediate analysis. However, microRNAs can simultaneously repress extensive suites of target genes (16). To provide experimental support for additional direct targets of miR-196 that have the potential to function in this developmental context, we identified the most highly up-regulated genes in our RNA-seq dataset that either contained a conserved binding site or were predicted to respond strongly to the miRNA (i.e., context+ score ≤ −0.2; Fig. 5A). For the top three evolutionarily conserved miR-196 target genes identified, we assessed whether regulation of their expression by miR-196 required direct binding to sites within their 3′ UTR. Using a luciferase-based reporter assay system in cell culture, miR-196 was shown to repress each of the target genes in a sequence-specific manner (Fig. 5B). Of particular interest within this set was the cell adhesion molecule (Prtg) involved in the ingestion of PSM progenitors (39) and an orphan nuclear receptor (Nr6a1) essential for somitogenesis in mice (40) and one of the very few genes that has been associated with variation of vertebral number (41). These experimentally supported miR-196 targets highlight important avenues for future investigation, not only with respect to axial patterning and elongation but also the many other divisions, and therefore total vertebral number, is promoted by trunk Hox proteins and antagonized by caudal Hox proteins (28–30). Our results place mir-196 activity at this critical junction, coordinating a reproducible trunk-to-tail Hox code transition. We suggest that a delay in Hox code transition could contribute to the formation of an additional vertebral element observed following genetic removal of miR-196 activity in mice. Such a role for miR-196 is likely to be broadly conserved across vertebrate species, as supported by regionalized vertebral expansion observed in miR-196 morphant zebrafish (24).

Wong et al. PNAS | Published online August 17, 2015 | E4889

Fig. 5. Identification of additional putative direct (non-Hox) miR-196 targets. (A) List of the most highly up-regulated genes and their associated fold changes in seven genotype comparisons that either (i) contain a conserved miR-196 binding site or (ii) are predicted to respond strongly to the miRNA. Genes with one or more conserved miR-196 target binding sites are indicated in green. (B) In vitro luciferase analysis confirms sequence-specific regulation of three experimentally supported target genes of miR-196. Renilla luciferase intensity values have been normalized to their respective Firefly values (RLU). Controls (WT 3′ UTR construct without miR-196b) were set to 1. MUT, mutated 3′ UTR construct destroying miR-196 binding site. Error bars represent SD. P values, Student t test: *P < 0.05, ***P < 0.0005, and ****P < 0.0001.

Fig. 6. Loss of miR-196 function alters signaling pathways known to control segmentation and axis elongation. Quantitative expression analysis of pathways known to control segmentation and axial extension in cells isolated from E9.5 mutant embryos, genotype comparisons are color-coded. Filled circles at the tips of fold changes represent a statistically significant change at q < 0.05.
observed following removal of all mature miRNAs in the mesoderm lineage (55), which in the latter case resulted in a caudal displacement of the hindlimb by three somites.

Finally, a coordinated temporal delay in the trunk-to-tail Hox code transition has been observed in mice null for Gdf11 (56), which as heterozygotes, bear striking phenotypic resemblance to 196a2+/−;196b−/− or miR-196 triple KO mouse embryos. We therefore specifically interrogated our RNA-seq data to assess the levels of Gdf11 and its direct downstream effector Isl1 (57). In 196a2+/−;196b−/− embryos, which exhibit 100% penetrant L-to-T transformation and sacral displacement, we observed a statistically significant reduction in Gdf11 and Isl1 levels by 15% (Dataset S1). As mentioned, this is likely to be an underestimate of the level of regulation, given the experimental strategy used. The requirement for Gdf11 in defining presacral vertebral number is dose dependent (56). The exact threshold requirement for Gdf11 signaling is not known, and it remains to be determined whether subtle down-regulation of Gdf11 contributes to phenotypic alterations observed in miR-196 mutant mice. Together, our transcriptome analyses revealed multiple developmental networks that require miR-196 activity for appropriate control of gene expression and suggest intriguing avenues for future experimental exploration.

**miR-196 Has the Potential to Modulate Wnt Signaling by Multiple Mechanisms.** Vertebral progenitors in the epiblast and tailbud are sensitive to the levels of Wnt signaling. Genetic removal of the Wnt3a ligand (58), or conversely, ectopic activation of Wnt3a in the epiblast (59), result in severe axis truncation posterior to the forelimb. Wnt3a expression has been shown to decrease as progenitor cells commit to a paraxial mesoderm fate (60, 61), and sustained Wnt activity disrupts somite formation (51) and somite polarity (59), dependent on timing and method of activation. These observations indicate that careful titration of Wnt levels is essential throughout the process of somite formation. Our data suggest that miR-196 activity is required in maintaining precise levels of Wnt activity (Fig. 6). Mechanistically, this could be achieved in at least two ways. First, miR-196 could directly target genes in the Wnt pathway. Specifically, the potent Wnt antagonist Dkk1 harbors a single predicted miR-196 site within its 3’ UTR, and Dkk1 expression was up-regulated following removal of miR-196 activity (Fig. 6). Using WISH, we confirmed increased expression of Dkk1 in 196a1−/−;196a2−/− embryos relative to 196a1+/−;196a2−/− (Fig. 7A; n = 2/2 per genotype). To test whether miR-196 can act directly to repress Dkk1, we used a luciferase-based reporter assay system in cell culture to show that indeed, miR-196 negatively regulates the Dkk1 3’ UTR in a sequence-specific manner (Fig. 7B). However, the repression in the reporter assay was more modest than that observed in vivo using RNA-seq (Fig. 6), and Dkk1 is not a conserved target of miR-196, suggesting that indirect regulation by miR-196 also plays a role. Second, miR-196 control over Wnt activity might work in part via Hox intermediates, which have the potential to either activate or repress Wnt signaling (29, 30). We recently showed using chick in vivo electroporation and imaging that the collarine activation of a subset of Hox9-13 posterior Hox genes within paraxial mesoderm progenitors translates into a graded increase in Wnt repression and a slowing down of axis elongation (30). One Hox gene that was found to significantly repress Wnt activity using this in vivo luciferase-based Wnt reporter assay was the miR-196 target Hoxa9. We therefore went on to test whether additional miR-196 Hox targets have the ability to repress Wnt activity in this context. We coelectroporated a Wnt/β-catenin reporter (BATLuc) and a CMV-Renilla construct in paraxial mesodermal progenitor cells, which express Hoxa9 and Hoxb1, and observed a significant reduction in luciferase activity in cells coexpressing the miR-196 reporter and Hoxa9 construct (Fig. 7C). Interestingly, these two target genes were regulated in a sequence-specific manner (Fig. 7D). However, the repression in the reporter assay was more modest than that observed in vivo using RNA-seq (Fig. 6), and Hoxa9 is not a conserved target of miR-196, suggesting that indirect regulation by miR-196 also plays a role. Second, miR-196 control over Wnt activity might work in part via Hox intermediates, which have the potential to either activate or repress Wnt signaling (29, 30).
Hox genes do not influence Wnt/β-catenin reporter activity in early chick paraxial mesoderm progenitors are the same Hox genes that show no indication of direct regulation by miR-196 in E9.5 mouse tissue (Fig. 4A). Together, these data demonstrate that miR-196 has the potential to directly and indirectly regulate the precise levels of Wnt activity in the developing embryo.

**Discussion**

Our work demonstrates the essential role for murine miR-196 in regulating vertebral identity across different levels of the body axis and reveals evolutionary conservation in the role of miR-196 in constraining total vertebral number. Importantly, our strategy has allowed us to comprehensively dissect paralog contribution to resultant phenotypes, allowing us to distinguish a patterning role for miR-196 from its role in modulating vertebral number. Moreover, we characterized the detailed molecular landscape controlled by miR-196 activity in the early embryo to show that miR-196 regulates, and therefore has the ability to integrate, multiple key signaling pathways to drive developmental processes.

**miR-196 Activity Is Essential for Vertebral Identity.** Despite the clear potential for functional redundancy between miR-196 paralogs (17), homeotic transformation of vertebral elements could be observed at low penetrance following removal of an individual miR-196 paralog (e.g., 196a2−/− or 196b−/− single mutants). With increasing loss of miR-196 family members (e.g., 196a2−/−;196b−/− double mutants), fully penetrant vertebral phenotypes were observed that were equivalent in severity to many single and compound Hox mutants (14, 62). Vertebral identity changes were observed at sites where loss-of-function phenotypes have previously been described for many of the direct Hox target genes (62), reinforcing the view that miR-196 acts within endogenous Hox domains rather than simply as a fail-safe mechanism to clear an anterior developmental program at more posterior locations (22). Paradoxically, the 196a2−/−;196b−/− or triple KO phenotypes are remarkably similar to either Hoxc8−/− or Hox8−/−;Hoxd8−/− skeletons, with eight ribs attached to the sternum, L1-to-T transformation, and a posterior displacement of the sacrum (62). However, with respect to number of sternal rib attachments and L1-to-T transformation, Hoxc8 loss-of-function and gain-of-function mutant mice exhibit identical phenotypes (31, 62). These data indicate that exquisite regulation of a quantitative Hox code is essential in defining vertebral identity at this axial location. Interestingly, deletion of Hoxb8 rescues many defects observed in Hoxc8-null mice, highlighting that there are aspects of a qualitative Hox code that we are yet to understand. Nonetheless, similar to Hoxc8, ectopic Hoxb8 expression results in supernumerary rib formation throughout the lumbar region (32), supporting the view that a collective up-regulation of direct Hox target genes drives homeotic alterations of the midthoracic to upper lumbar region in miR-196 mutant mice.

A shift in the position of the sacrum observed in miR-196 mutant embryos was not easily reconcilable with the function of miR-196 in directly repressing trunk Hox target genes (31, 32). However, we show that in addition to direct Hox gene regulation, miR-196 indirectly regulates the expression levels or temporal activation of many caudal Hox genes, including those whose protein products are known to control positioning of the sacrum, such as Hoxa10, Hoxd10, and Hoxd11 (13, 14, 63, 64). The mechanisms leading to a delay in posterior Hox gene activation in miR-196 mutant mice are currently unknown. A similar coordinated temporal shift in the trunk-to-tail Hox code has been demonstrated in Gdf11−/− mice (56), which show conservation in the types of vertebral transformations we observe here in miR-196 mutant embryos. In this context, Gdf11 appears to work via retinoic acid signaling (57, 65), and whether altered Gdf11 and retinoic acid signaling contribute to miR-196 phenotypic alterations remains to be tested.

**miR-196 Activity Constrains Total Vertebral Number.** Total vertebral number of a given species is highly reproducible, and mutations that extend the vertebral column of model organisms are very rare. Among vertebrate species, however, great diversity in vertebral number has arisen. Cross-species comparison (8) or direct genetic perturbation (25, 27) demonstrates that the periodicity of segmentation clock oscillation relative to the rate of PSM growth is the central parameter in defining vertebral number. It remains to be determined how an additional vertebral element seen here in miR-196 mutant mice, or in miR-196 morphant zebrafish (24), are generated at a cellular level (i.e., does the clock tick faster or does it tick at the same rate for longer). Our analysis does, however, reveal molecular alterations in miR-196 mutant embryos that have the potential to affect vertebral number.

First, altered expression of Notch, Wnt, and Fgf pathways could alter the periodicity of segment formation (1). However, diminished Wnt and Fgf would be predicted to increase somite size (27, 66–68), which if axis elongation was unaltered, would lead to a reduction in vertebral number. Further work is required to clarify any functional role for miR-196 in the molecular networks coordinating segmentation.

Second, we showed that miR-196 activity can modulate the expression levels of many Hox genes, either directly or indirectly. It is well documented that Hox genes control mesodermal ingress, thus regulating cell injection into the PSM (30, 69). The rate of PSM growth is not uniform along the anterior-posterior axis (8), with a switch to PSM shortening occurring at about the trunk-to-tail transition in most amniotes. This switch correlates with activation of a posterior Hox code (Hox9 onward), and a subset of posterior Hox genes slows axis elongation by controlling the ingress of PM progenitors via Wnt repression (30). We show here that the ability to repress Wnt signaling is not exclusive to posterior Hox genes but that Hox7/8 paralogs also downregulate Wnt signaling in the chick epiblast. The repression of Wnt by posterior Hox genes as a means to slow down and terminate axis elongation (29, 30) is consistent with the known function of Wnt3a in driving axis elongation (58). The repression of Wnt by trunk Hox genes is less intuitive because, at least on a Cdx2/4 mutant background, ectopic Hoxb8 has been shown to maintain Wnt expression in the tailbud (29). However, the importance of precise Wnt levels in the early steps of axis formation, and of cellular context, are beginning to be appreciated (59, 70) and may underlie such discrepancies. Exactly how miR-196 activity integrates within the genetic networks controlling PM progenitor maintenance and differentiation remains to be fully elucidated. At E9.5, this progenitor population represents quite a minor fraction of miR-196–expressing cells. Therefore, the extent to which molecular alterations we identified in miR-196 mutants, including reduced Wnt/Fgf signaling, reflect an altered PM progenitor molecular status requires clarification at a cellular resolution.

Although a heterochronic shift in the trunk-to-tail Hox code transition could be predicted to vary vertebral number, morphological evidence for this has been scarce. Analysis of total vertebral number in Gdf11−/− mice, which exhibit a dramatic heterochronic shift in Hox code, is hampered by caudal truncation (56). Although ectopic trunk Hox expression (Hoxa5 and Hoxb8) has the ability to rescue axis truncation defects of a genetically engineered mutant (29), they do not appear to increase vertebral number on a WT background (29, 31). This lack of phenotype is possibly due to the fact that posterior prevalence still holds; caudal Hox genes and miR-196 would be expressed at the usual time and place to regulate and terminate axis elongation. In the case of miR-196 KO’s, the cumulative effect on both trunk and caudal Hox gene expression could permit continued maintenance of progenitor divisions while delaying commencement of axis elongation slow down, resulting in increased vertebral number.

Together, our results highlight an essential requirement for miR-196 activity in reinforcing a timely trunk-to-tail Hox code
transition and reproducibility of axial formulae. Given the ancestral role of Hox activity in species that use a posterior growth zone (71), and the recurrent acquisition of miRNAs within the Hox clusters across metazoan taxa (15, 17, 72, 73), variation in Hox-miRNA interactions may represent an important mechanism for the evolution of animal body plans.

Materials and Methods

miR-196a1\textsuperscript{exp} and miR-196a2\textsuperscript{exp} Knock-In Construction. A 72- bp (miR-196a1) or 52-bp (miR-196a2) genomic fragment encompassing each mature miRNA sequence was replaced with a cassette containing eGFP fused to the rabbit β-globin 3' UTR followed by flipase recognition target (FRT)-flanked PGKem7-Neomycin. A Kozak sequence was inserted upstream of the eGFP start codon. Targeting constructs were generated using a 129sv sequence and electroporation into J1 embryonic stem cells. Correctly targeted ES cells were identified and used to generate germ-line transmitting knock-in lines. Before analysis, the Neomycin selection cassette was removed by crossing to a ubiquitous FlPe-deleter mouse line. Resulting lines were bred onto a C57BL/6J background and confirmed as isoegenic by SNP genotyping.

miR-196a1\textsuperscript{−/−}, miR-196a2\textsuperscript{−/−}, and miR-196b\textsuperscript{−/−} Generation. Previously targeted ES cells at each of the three miR-196 loci have been generated (74). Correctly targeted JMB3 ES cells were reconfirmed by Southern blot and used to generate germ-line transmitting KO lines. Before analysis, the puDeltaTK selection cassette was removed by crossing to a ubiquitous Cre-deleter mouse. Resulting lines are on a mixed C57BL/6J and C57BL/6N background.

Mouse Skeletal Preparation and Analysis. Skeletal preparation was performed on E18.5 embryos or p6 postnatal pups as previously described (75).

In Situ Hybridization. Whole mount in situ hybridization was performed as previously described (76).

FACS Sorting and RNA-seq Sample Preparation. Freshly dissected E9.5 embryos were dissociated in 0.25% trypsin/2% (vol/vol) cherk serum, neutralized in DMEM + 10% (vol/vol) FBS, and washed into PBS + 2% (vol/vol) FBS for FACS sorting into J1 embryonic stem cells. Correctly targeted ES cells were buffer (Qiagen) and RNA isolated using RNeasy with added on-column DNase treatment (Qiagen). RNA quality was assessed using a Bioanalyzer, and 200 ng per individual embryo was used as input for RNA-seq library generation (unstranded illumina TruSeq Kit). Libraries were multiplexed and sequenced using an Illumina HiSeq. 2000 instrument, generating 50-bp single end reads.

RNA-seq and Category Enrichment Analysis. Quantification of the transcriptome using RNA-seq data was performed as previously described (77). Raw reads were aligned to the latest build of the mouse genome (mm10) using STAR v. 2.3.1n (options-outFilterType BySlout-outFilterMultimapScoreRange 0-reads-MatesLengthsInEQ-EqualFilterIntronMotifs RemoveNoncanonical Unannotated-dipladAdapter SequencesTGTAGATCCGCTTCTGCTGTTG-outSAMStrandField intronMotif-outStd SAM) (78). Considering all replicates of a particular genotypt, differential expression statistics were computed between genotypes of interest using cuffdiff v. 2.1.1 (options outStd SAM) (78). Considering all replicates of a particular genotypt, differential expression statistics were computed between genotypes of interest using cuffdiff v. 2.1.1 (options-library-type fr-unstranded -c 100 -b mm10.fa -u -outSAMstrandField intronMotif-outStd SAM) (78). Considering all replicates of a particular genotypt, differential expression statistics were computed between genotypes of interest using cuffdiff v. 2.1.1 (options-library-type fr-unstranded -c 100 -b mm10.fa -u -out-max-bundle-fraggs 1000000000) (79), using mouse transcript models of protein-coding genes annotated in Ensembl release 72. Before all subsequent analyses, we filtered away genes annotated by cuffdiff as "NOTEST" in all genotypes, indicating the genes were too lowly expressed to accurately quantify their abundances. To evaluate functional gene categories that were statistically enriched, we loaded differentially expressed genes (i.e., genes with Q < 0.05) into the Core Analysis function of Ingenuity Pathway Analysis software (Ingenuity Systems), testing gene categories related to development and function. All P values reported from this analysis were adjusted using the Benjamini-Hochberg method to control the false discovery rate.

miRNA Target Analysis. To identify predicted miRNA targets, the 3' UTR sequences of protein-coding genes were searched to identify 6mer, 7mer-A1, 7mer-miR, and 8mer-miRNA binding sites cognate to the miR-196 seed (35, 80). A context+ score was computed for each target site within a given 3' UTR, and scores were summed to produce a total context+ score for each gene, which was used for all miRNA-related analyses (35). TargetScanMouse 6.2 was further used to assess target site conservation or to include predicted miR-196 targets containing noncanonical 3' compensatory sites, such as in the case of Hoxb8 (81).

Permutation Test for Significance Testing. A permutation test was devised to evaluate the significance of differences in vertebral number. Briefly, given two groups of count-based data of size n and m, we randomly partitioned the counts (without replacement) from the union of the two groups to generate 100,000 pairs of data, again of size n and m. To compute an empirical one-sided P value, we then computed the proportion of pairs that satisfied the condition that the difference in the means of each pair exceeded the difference in the means of the original two groups.

In Vitro Luciferase Assay. A 3' UTR sequence (300–700 nucleotides) of protein-coding genes of interest were commercially synthesized and cloned into pScheck2 vector. For each, a mutant version containing four nucleotide substitutions within the miR-196 seed sequence was generated. Constructs were transfected into NIH 3T3 cells with or without 25 pmol mmu-miR-196b duplex. Transfection (Lipofectamine2000; Life Technologies) and luciferase analysis (Dual Luciferase Reporter Assay System; Promega) were performed as per the manufacturer's instructions.

Chick Electroporation and in Vivo BatLuc Reporter Analysis. Chicken embryos were harvested at Hamburger and Hamilton stage 5 (82) and electroporated ex ovo as previously described (30), with a DNA mix containing BATLuc (1 μg/μl final), CMV-Renilla (Promega; used as a control to normalize the differences of electroporation intensity between embryos; 0.2 μg/μl final), a control pCAGGS-Venus vector (gift from K. Hadjantonakis, Memorial Sloan Kettering Cancer Center, NY), or a Hox gene of interest (Hoxb7, a5, a7, b7, b8, or d8; cloned in pCAGGS-IRES-Venus (5 μg/μl final). Electroporated embryos were cultured in a humidified incubator at 38 °C for 20 h. Embryos were analyzed using a fluorescent microscope, and only embryos showing restricted expression of Venus in the paraxial mesoderm were selected (90–100% of the electroporated embryos) for luciferase assay (between three and five embryos for each condition). The posterior region (from somite 1 to tailbud) of the selected embryos was dissected and lysed in passive lysis buffer (Promega) for 15 min at room temperature. Lysates were then distributed in a 96-well plate, and luciferase assays were performed using a Centro LB 960 luminometer (Berthold Technology) and the dual luciferase kit (Promega) following manufacturer’s instructions. Raw intensity values for Firefly luciferase signal were normalized with corresponding Renilla luciferase values (RLUs), and the control experiment was set to 1.

Acknowledgments. We thank Xin Sun and Denis Duboule for supplying in situ probes and A. Dobin for help in understanding RNA-seq mapping parameters. We thank Allan Bradley for providing three KO mouse ES cell lines used in these studies. We thank Christophe Marcell, Erin Hornstein, and Jan Manent for critical reading of the manuscript. This work was supported by a National Science Foundation Graduate Research fellowship (to V.A.), Wellcome Trust Grant 098051 (to support miRNA KO reagent production by H.M.P.), National Institutes of Health Grant GM067031 (to D.P.B.), National Institutes of Health Grant R37 HD02443-19 (to C.J.T.), and National Health and Medical Research Council Project Grant APP1031792 (to E.M.). E.M. thanks Bioplatforms Australia for support. D.P.B. is a Howard Hughes Medical Institute Investigator. The Australian Regenerative Medicine Institute is supported by grants from the State Government of Victoria and the Australian Government.


