

RESEARCH ARTICLE

# Hotair Is Dispensable for Mouse Development

Ana Rita Amândio<sup>1</sup>, Anamaria Necsulea<sup>1</sup>, Elisabeth Joye<sup>1</sup>, Bénédicte Mascrez<sup>2</sup>, Denis Duboule<sup>1,2\*</sup>

**1** School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland,

**2** Department of Genetics and Evolution, University of Geneva, Geneva, Switzerland

\* [denis.duboule@epfl.ch](mailto:denis.duboule@epfl.ch), [denis.duboule@unige.ch](mailto:denis.duboule@unige.ch)



CrossMark  
click for updates

 OPEN ACCESS

**Citation:** Amândio AR, Necsulea A, Joye E, Mascrez B, Duboule D (2016) *Hotair* Is Dispensable for Mouse Development. *PLoS Genet* 12(12): e1006232. doi:10.1371/journal.pgen.1006232

**Editor:** Gregory S. Barsh, Stanford University School of Medicine, UNITED STATES

**Received:** April 18, 2016

**Accepted:** July 8, 2016

**Published:** December 15, 2016

**Copyright:** © 2016 Amândio et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Raw and processed RNA-seq data are available in GEO, accession number GSE79028.

**Funding:** This work was supported by institutional funding from the EPFL (<http://www.epfl.ch>) and the University of Geneva (<http://www.unige.ch>) as well as by grants from the Swiss National Research Fund ([www.snf.ch](http://www.snf.ch)) (No. 310030B\_138662) and the European Research Council (<https://erc.europa.eu>) (ERC SystemsHox.ch, No: 232790 and ERC RegulHox, No: 667156) (to DD). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

Despite the crucial importance of *Hox* genes functions during animal development, the mechanisms that control their transcription in time and space are not yet fully understood. In this context, it was proposed that *Hotair*, a lncRNA transcribed from within the *HoxC* cluster regulates *Hoxd* gene expression *in trans*, through the targeting of Polycomb and consecutive transcriptional repression. This activity was recently supported by the skeletal phenotype of mice lacking *Hotair* function. However, other loss of function alleles at this locus did not elicit the same effects. Here, we re-analyze the molecular and phenotypic consequences of deleting the *Hotair* locus *in vivo*. In contrast with previous findings, we show that deleting *Hotair* has no detectable effect on *Hoxd* genes expression *in vivo*. In addition, we were unable to observe any significant morphological alteration in mice lacking the *Hotair* transcript. However, we find a subtle impact of deleting the *Hotair* locus upon the expression of the neighboring *Hoxc11* and *Hoxc12* genes *in cis*. Our results do not support any substantial role for *Hotair* during mammalian development *in vivo*. Instead, they argue in favor of a DNA-dependent effect of the *Hotair* deletion upon the transcriptional landscape *in cis*.

## Author Summary

During mammalian embryonic development, *Hox* genes must be tightly regulated. It was proposed earlier that part of this regulation relies upon *Hotair*, a long non-coding RNA that recruits repressive protein complexes onto the *HoxD* gene cluster to keep these genes silent before they become activated. A genetic deletion of *Hotair* in mice induced homeotic transformations, thus supporting this hypothesis. However, other alleles involving this locus gave controversial results and hence we re-assessed the effect of the full deletion of *Hotair in vivo*. In our genetic background and using our analytical conditions, we could not confirm the reported morphological alterations, nor could we detect any mis-regulation of *Hoxd* genes in those fetal tissues where *Hotair* is detected in control animals. However, the genomic deletion induces the mis-regulation *in-cis* of the neighboring *Hoxc11* and *Hoxc12* genes, a side-effect which may underlie a weakly penetrant alteration observed in the shape of some tail vertebrae.

**Competing Interests:** The authors have declared that no competing interests exist.

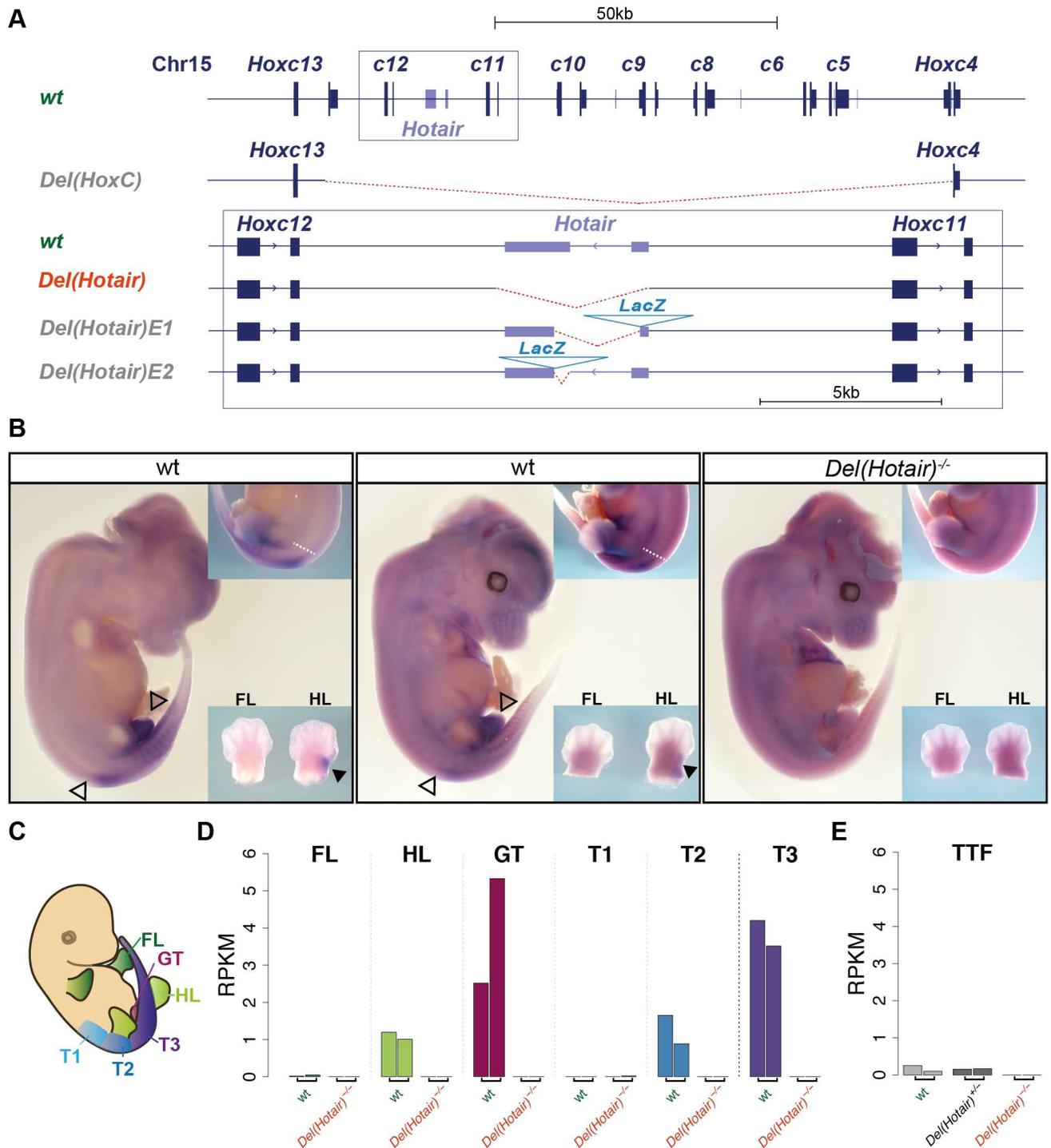
## Introduction

*Hox* genes encode transcription factors with crucial roles in the specification of regional identities along the body axes during development. Mutations affecting specific *Hox* genes typically lead to homeotic transformations, whereby a particular body part is transformed into the identity of another one [1–4]. In mammals, following the two rounds of genome duplication that occurred at the basis of the vertebrate lineage (see [5]), four distinct clusters of *Hox* genes are found (*HoxA* to *HoxD*) (ref. in [6]). During development, *Hox* genes are transcriptionally activated in a precise temporal and spatial sequence, which follows their chromosomal order [7,8]. These collinear patterns of transcription are regulated at multiple levels and studies focusing on the *HoxA*, *HoxB* and *HoxD* loci have revealed the importance of intricate combinations of local and long-range *cis*-regulatory elements. Also, studies using micro-dissected embryonic material have shown that the transcriptional activation of these genes, in different ontogenetic contexts, is accompanied by major changes in both the epigenetic modifications of the surrounding chromatin and its 3D spatial organization [9–11].

Long non-coding RNAs (lncRNAs) have been proposed to represent yet another layer of regulatory control at these important developmental loci (e.g. [12–15]). Increasing evidence indeed suggests that lncRNAs can act as regulators of gene expression, for example by interacting with transcription factors and chromatin modifiers to modulate transcription during development [16]. Several lncRNAs associated with the mammalian *Hox* clusters have been identified, amongst which *Hotair* (*Hox* transcript antisense intergenic RNA), a lncRNA transcribed from the intergenic region between *Hoxc11* and *Hoxc12* within the *HoxC* cluster and the founding member of this new class of RNAs. *Hotair* was proposed to help repress some 5'-located (posterior) *Hoxd* genes in *trans*, through its association with chromatin modification complexes such as PRC2, LSD1 and CoREST/REST [13,17]. Accordingly, *Hotair* would recruit or enrich this part of the *HoxD* cluster with *Polycomb* (*Pc*) complex, thus contributing to its repressed state before transcription starts. This proposal was substantiated by the knockdown of *Hotair* in human fibroblasts, which led to a decreased binding of *Pc* repressive complexes in the *HoxD* cluster and to a concurrent increase in *Hoxd* genes expression [13].

This important function for a lncRNA in cultured human fibroblasts was however not supported by the analysis of a mouse line carrying a targeted deletion of the entire *HoxC* cluster [18], i.e. including the mouse *Hotair* lncRNA. This deletion showed little effect *in vivo*, with no alteration of *Hoxd* genes expression. Also, the presence and enrichment of H3K27me3 repressive chromatin marks at the *HoxD* locus was not dramatically modified [19]. This lack of effect was tentatively explained by the concomitant *in-cis* deletion of all *Hoxc* genes, which may have masked or compensated a potential alteration caused by the absence of *Hotair* alone [20]. To alleviate this problem, three alleles were recently produced where the *Hotair* transcript was specifically targeted (Fig 1). The first allele is a targeted deletion of the two major exons of *Hotair*. Mice carrying this deletion were reported to display a malformation of the wrist and homeotic transformations of the spine, either from six lumbar vertebrae (L6) to a L5 vertebral formula, or within the post-sacral region [20]. These phenotypes were associated with a depression of *Hoxd* genes and of a set of imprinted genes by modulation of their chromatin state [20]. These effects were scored in the absence of any change in the transcription of the neighboring *Hoxc11* and *Hoxc12* genes, supporting a function of *Hotair* in *trans* [20].

Two additional *Hotair* deletion mutant alleles combined with *LacZ* reporter knock-in were recently reported by Lai and colleagues [21]. The first allele deleted nearly the entire *Hotair* sequence and the second one comprised a smaller deletion starting in the second exon [21]. In both cases, while a subtle alteration of the 4th caudal vertebra was scored, the wrist and the spine appeared normally formed, without any sign of the lumbar homeotic transformation



**Fig 1. *Hotair* expression in vivo.** (A) Schematic representation of the wild type *Hotair* locus and the various *Hotair* deletion alleles. The deleted DNA is in red. The *HoxC* allele is from [18] and the shorter deletions in the box from [20] and [21]. (B) Whole mount in situ hybridization (WISH) of *Hotair* RNAs on E12.5 wild type CD1 (left) and CBA/C57/B6 (center) control embryos and of a *Del(Hotair)<sup>-/-</sup>* mouse embryo (right, n = 3). No signal was detected in *Del(Hotair)<sup>-/-</sup>* embryos, demonstrating the specificity of the probe. *Hotair* is expressed with a posterior restriction (white dashed line), resembling the transcript distribution of either a *Hox11* or a *Hox12* gene. Black arrowheads indicate expression domain of *Hotair* in the hindlimbs, hollow arrowheads indicate the limit of *Hotair* expression in the trunk and in the genital tubercle. The common artifact signal in the cerebral vesicles results from incomplete opening of these vesicles and subsequent probe trapping. (C) Schematic representation of the dissection patterns for RNA-seq. These dissections involved forelimbs (FL, dark green), hindlimbs (HL, green) and the genital tubercle (GT, magenta), as well as three trunk sections corresponding to the lumbar/sacral (T1, light blue); sacro/caudal (T2, blue) and caudal (T3, purple)

regions. (D) Quantification of *Hotair* expression by RNA-seq (normalized RPKM values). (E) Quantification of *Hotair* expression (normalized RPKM values) in tail tip fibroblasts (TTF), using data from [20].

doi:10.1371/journal.pgen.1006232.g001

and wrist alterations previously reported for the deletion of both exons [20]. Due to our long-lasting interest in the transcriptional regulation of *Hoxd* genes during development (e.g. [22]), we addressed these apparently conflicting results by re-assessing the effects of deleting the *Hotair* lncRNA during early mouse development, using the largest deletion allele previously described [20]. In agreement with earlier and more recent results [19,21], we find that the deletion of *Hotair* has no substantial effect, neither on wrist morphology, nor on the vertebral formula at the lumbo-sacral level. In addition, transcriptome analyses reveal that the absence of *Hotair* does not impact upon *Hoxd* genes expression *in trans*, in any of the embryonic tissues analyzed. In contrast, we observe subtle yet significant changes in the expression of the neighboring *Hoxc11* and *Hoxc12* genes in the mutant mice, indicating an *in-cis* impact of modifying the genomic locus. Taken together, our results strongly suggest that the *Hotair* lncRNA has little effect—if any—on mouse embryonic development.

## Results

### *Hotair* expression *in vivo*

We extended the analysis of a mouse strain that includes a deletion of the two major *Hotair* exons (Fig 1A)[20]. Even though we concluded that this mutation is primarily an allele of the *HoxC* cluster (see below) and hence that it should be referred to as *HoxC<sup>Del(Hotair)</sup>*, we shall refer to it as *Del(Hotair)* throughout this study for sake of simplicity. We first confirmed the expression pattern of *Hotair in vivo* by whole mount *in situ* hybridization (WISH) using wild type mice of two distinct genetic backgrounds (CD1 and CBA/C57/B6) as well as *Del(Hotair)<sup>-/-</sup>* mouse embryos at embryonic day 12.5 (E12.5)(Fig 1B). Staining of *Del(Hotair)<sup>-/-</sup>* embryos confirmed the specificity of the *Hotair* probe as no signal was detected in these embryos (Fig 1B). In contrast, wild type embryos of both genetic backgrounds showed the presence of *Hotair* transcripts in the genital tubercle, the proximal part of the hindlimbs and in the caudal part of the embryo (Fig 1B), confirming previously published data [19] and consistent with the *LacZ* staining reported for the *Hotair* knocked-in allele [21]. In both cases, staining was observed just posterior to the lumbar region and was not scored in developing forelimbs [21].

We quantified the expression levels of *Hotair* with high coverage RNA-sequencing (RNA-seq) (Materials and Methods, S1 Table). Based on the spatial expression pattern of *Hotair* as determined with WISH and on the skeletal phenotypes reported in mice by Li et al. [20], we micro-dissected both *Del(Hotair)<sup>-/-</sup>* and wild type E12.5 embryos into six distinct parts for comparative RNA-seq analyses (Fig 1C). We thus separately collected the forelimbs (FL), the hindlimbs (HL), the genital tubercle (GT), a piece of trunk corresponding to the lumbo-sacral region (T1); a piece of trunk corresponding to the sacro-caudal region (T2) and finally, a piece of trunk corresponding to the developing caudal region (T3, Fig 1C). As expected from the WISH experiments, *Hotair* transcripts were scored in the hindlimbs, the genital tubercle and the trunk samples T2 and T3. The highest steady-state levels of *Hotair* RNAs were detected in the GT and the T3 embryonic tissues (Fig 1D, S1 Dataset). As a control, *Hotair* transcripts were not detected in any tissues derived from homozygous *Del(Hotair)<sup>-/-</sup>* mutant embryos (Fig 1D). To better compare this dataset with published results, we analyzed in parallel the RNA-seq data obtained from primary tail tip fibroblast (TTF), derived from both wild type and *Del(Hotair)<sup>-/-</sup>* mice [20]. This analysis revealed that the expression level of *Hotair* in control TTF

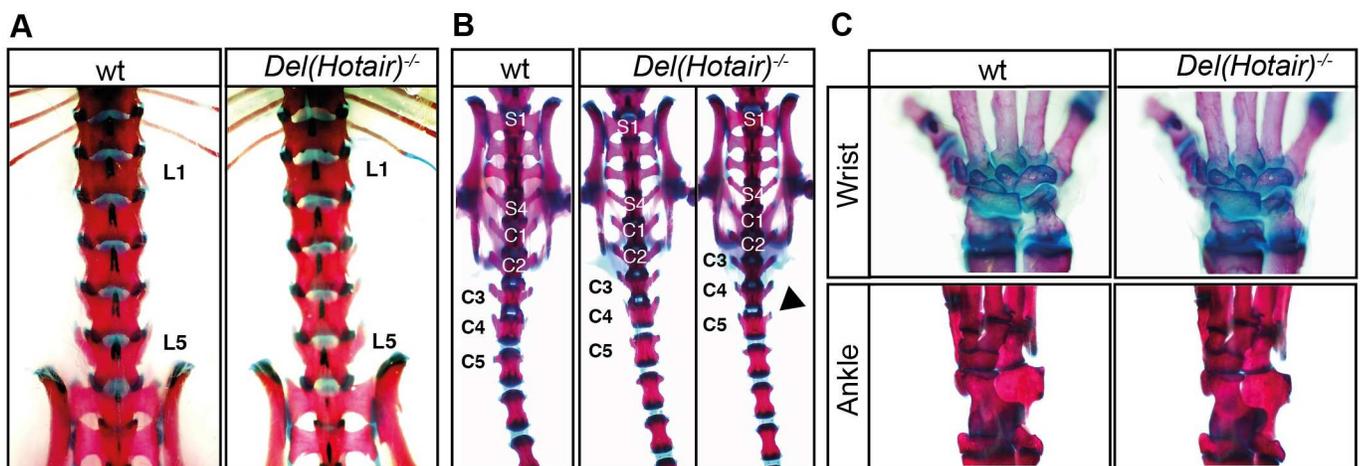
was very low when compared to its expression levels in the GT or the posterior T3 trunk sample (Fig 1E, S2 Dataset).

### Phenotype of mice lacking the *Hotair* lncRNA

*Hotair* was reported to be important for both the proper establishment of the mouse vertebral column and for the formation of the forelimb mesopodial articulation: the wrist [20,21]. To confirm this phenotypic effect, we inter-crossed *Del(Hotair)<sup>+/-</sup>* heterozygous mice and examined the skeletons of F1 animals at postnatal day 22 (P22). We investigated in particular the three reported sites of observed alterations in mutant *Del(Hotair)<sup>-/-</sup>* mice [20,21]. We first searched for potential differences in vertebral formulae, as it was reported that 58% of *Del(Hotair)<sup>-/-</sup>* mice had five lumbar vertebrae, while 100% of wild type CBA/C57/BL6 mice had six lumbar vertebrae [20].

All our mutant alleles at *Hox* loci (see e.g. [23]) are systematically backcrossed onto mixed (B6xCBA)F1 animals to maintain heterogeneous but similar backgrounds when comparing experimental crosses. After bringing the *Hotair* mutant mice [20] onto this genetic background for some generations, we found that 80% of *Del(Hotair)<sup>-/-</sup>* mice displayed five lumbar vertebrae, similar to wild type littermates (Chi-square test, *p*-value 0.97, Fig 2A, Table 1). In both wild type and homozygous mutant animals, the L6 formula was sporadically scored, as well as the mixed L6/S1 vertebral type, often observed in our stocks. Despite a limited number of specimens observed, but together with the fact that we were unable to detect specific transcripts in the embryonic trunk at this vertebral level by two independent methods, we conclude that this lncRNA is very unlikely to have a function in the organization of this very flexible morphological boundary (see discussion).

We next analyzed the morphology of caudal vertebrae in the post-sacral region. Previous analyses had concluded that mice with *Hotair* deletions had longer lateral processes on the fourth vertebra when compared to wild type animals, with full penetrance. In our case, we observed that three out of ten *Del(Hotair)<sup>-/-</sup>* mutant mice had longer processes on the fifth caudal vertebra, compared to wild type (Fig 2B). This may indeed correspond to a very subtle



**Fig 2. *Hotair* deletion has little impact, if any, on skeletal morphology.** Alizarin Red and Alcian Blue skeletal staining of wild type and *Del(Hotair)<sup>-/-</sup>* mice. (A) Lumbar region of wild type (left) and *Del(Hotair)<sup>-/-</sup>* (right). In our (B6xCBA) background, both control and mutant animals have five lumbar vertebrae (L5), with an equally low incidence of L6 (see Table 1). (B) The sacro-caudal region of wild type (left) and *Del(Hotair)<sup>-/-</sup>* (right) animals, with the black arrowhead pointing to a moderate gain of lateral protrusion in mutant caudal vertebra 5 (C5), usually not observed in control animals. (C) Normal wrist and ankle bones in both wild type (left) and *Del(Hotair)<sup>-/-</sup>* (right) animals. The number and organization of mesopodial bones remained unchanged in the mutant condition.

doi:10.1371/journal.pgen.1006232.g002

**Table 1. Phenotypic analysis of wild type and *Del(Hotair)*<sup>-/-</sup> skeletons at post-natal day 22 (P22).**

|                  | wt<br>(n = 11) | <i>Del(Hotair)</i> <sup>+/-</sup><br>(n = 11) | <i>Del(Hotair)</i> <sup>-/-</sup><br>(n = 10) |
|------------------|----------------|---|---|
| Lumbar vertebrae |                |   |   |
| L6               | 1              | 1   | 1   |
| L6/S1*           | 2              | 1   | 1   |
| L5               | 8              | 9   | 8   |
| Wrist            |                |   |   |
| normal           | 11             | 11  | 10  |
| ill-formed       | 0              | 0   | 0   |
| Ankle            |                |   |   |
| normal           | 11             | 11  | 10  |
| ill-formed       | 0              | 0   | 0   |

\* Mixed identity

Various control and mutant specimen were scored for having either five lumbar vertebrae (L5), six (L6) or a mixed L5-6/S1 vertebra. The mesopodial bones (wrist and ankle) were found normal in shape and number in all cases.

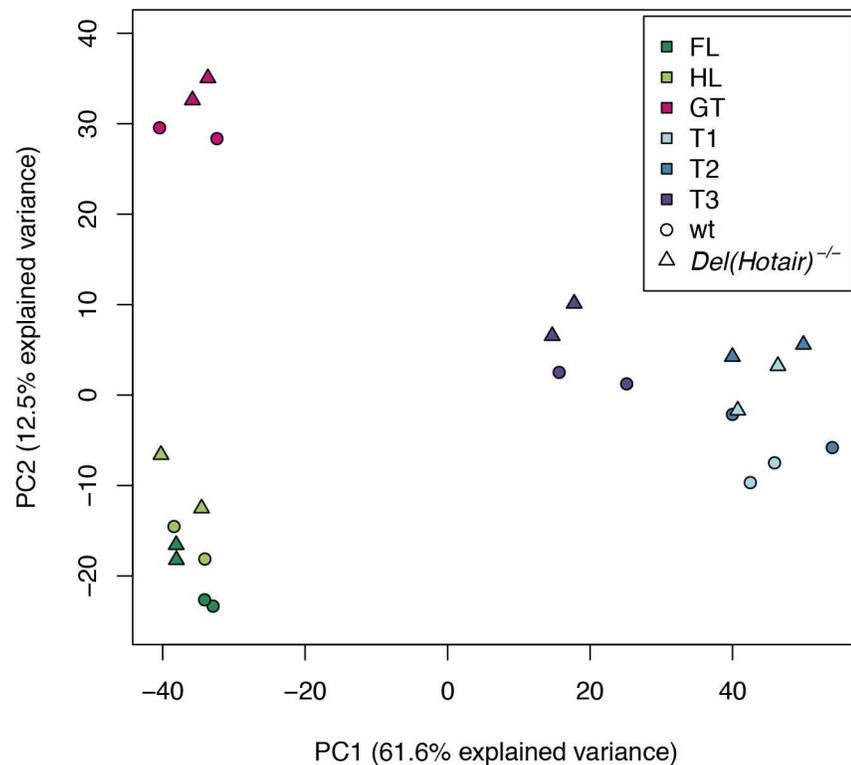
doi:10.1371/journal.pgen.1006232.t001

morphological alteration in this region of the caudal spine, although the penetrance of this light phenotype is not 100%. Unlike the lumbo-sacral and wrist alterations [20], this particular tail vertebral morphology was also scored by Lai and colleagues when analyzing another mutant allele of *Hotair* [21].

Finally, and even though we were unable to detect any *Hotair* transcripts in the forelimbs of E12.5 mice embryos, unlike for hindlimbs (Fig 1B), we carefully examined both forelimb and hindlimb skeletons of wild type and *Del(Hotair)*<sup>-/-</sup> mutant mice. We did not detect any alteration in limb morphology (Fig 2C and Table 1), in particular in the anatomy of the wrist, where malformations due to the loss of *Hotair* had been previously reported (Fig 2C). The same conclusion was reached concerning the hindlimbs, even though *Hotair* transcripts could clearly be scored in the proximal part. Altogether, we could not reproduce the reported phenotypic effects of *Hotair* deletion at two sites, the wrist and the lumbo-sacral region, where we were also unable to detect any *Hotair* transcripts. Regarding tail vertebrae, a slight effect could indeed be observed, poorly penetrant and likely dependent on the genetic background (see below).

### Transcription profiles of *Del(Hotair)*<sup>-/-</sup> mutant embryonic tissues

To more globally evaluate the effect of *Hotair* deletion upon developmental gene regulation, we performed a principal component analysis (PCA) using the expression levels of all autosomal protein-coding genes detected in our RNA-seq experiments (Materials and Methods). We observed a good separation of the data according to tissue type, although the T1 and T2 samples clustered together (Fig 3). Principal component 1 (PC1), which explained 61.6% of the total gene expression variance, separated the trunk segments (T1, T2 and T3) from the other embryonic tissues. Likewise, the differences between GT, HL and FL were resolved along PC2, which accounted for 12.5% of the total variance (Fig 3). Part of the variance was also explained by the genotypes and we observed that wild type and *Del(Hotair)*<sup>-/-</sup> samples were separated along PC2 (Fig 3). Of note, the same separation between wild type and *Del(Hotair)*<sup>-/-</sup> on PC2 was observed in all tissues, even in T1 and FL where *Hotair* is not expressed (Fig 3). In agreement with the results from the clustering between samples, we observed high expression level



**Fig 3. Overview of gene expression patterns in wild type and *Del(Hotair)<sup>-/-</sup>* embryonic tissues.** First factorial map for the principal component analysis (PCA) of gene expression levels. Tissues are color-coded (upper right corner) and the genotypes are indicated by either a circle (wild type) or a triangle (*Del(Hotair)<sup>-/-</sup>*). The numbers in parentheses indicate the proportion of the variance explained by PC1 or by PC2.

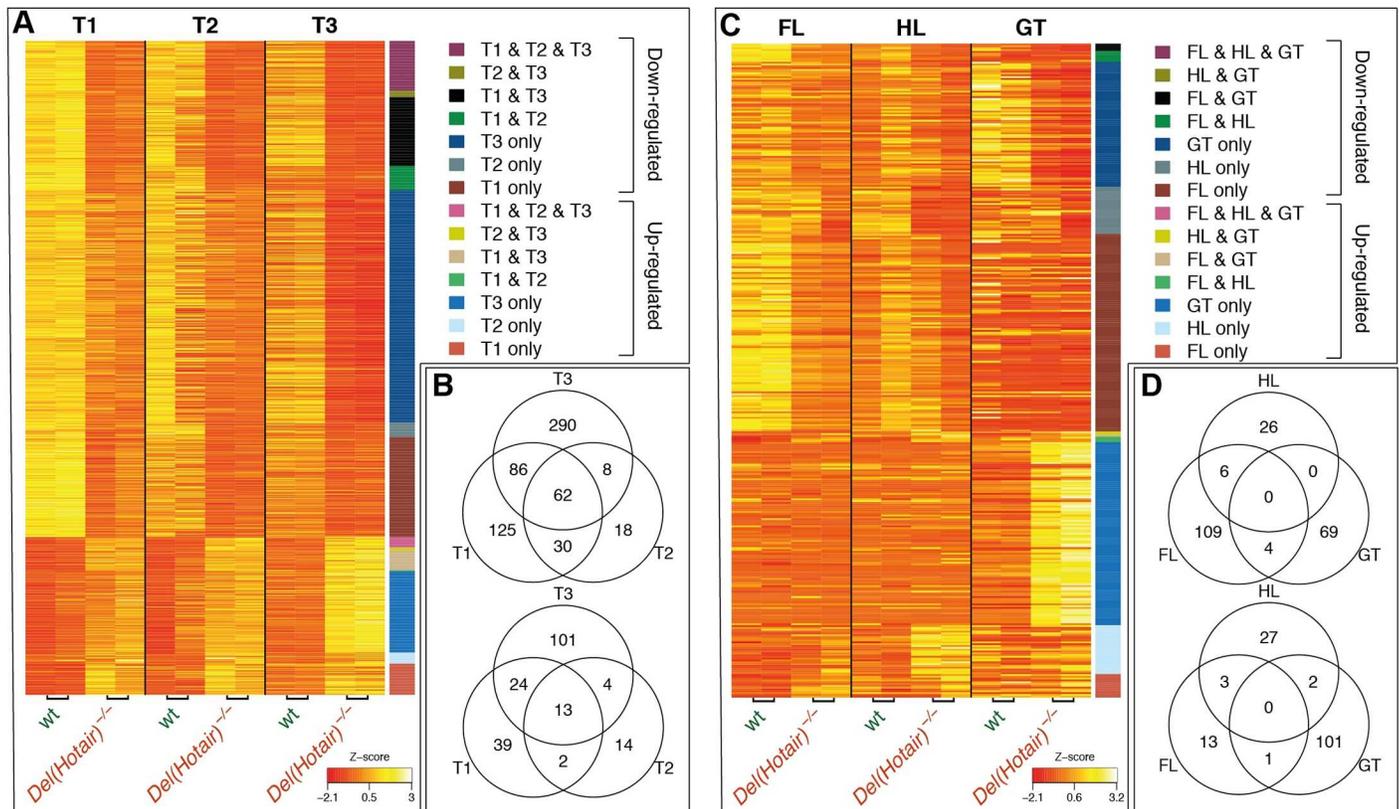
doi:10.1371/journal.pgen.1006232.g003

correlations among biological replicates for all tissues (S1 Fig). Furthermore, gene expression clustering based on pairwise Euclidean distances between samples (see [Materials and Methods](#)) showed a clear separation between four different groups: the limbs (FL and HL), the GT, the T3 trunk segment and the remaining T1 and T2 trunk segments (S2 Fig). Using this method, we observed a separation between genotypes only when the T3 sample was considered.

Altogether, these results point to the reproducibility of replicates and illustrate the good separation between tissues, with the exception of the T1 and T2 trunk samples. The high similarity in global gene expression between T1 and T2 likely reflects the spatial proximity of these two tissues, even though we cannot exclude some variation in the positioning of the precise T1/T2 boundary during dissection, which is a challenging task in such young embryos.

### Expression analysis of wild type *versus Del(Hotair)<sup>-/-</sup>* embryonic tissues

Since *Hotair* was proposed to act as a repressor of gene expression in cultured fibroblasts [13,20], we conducted tissue-specific differential gene expression analyses between wild type and *Del(Hotair)<sup>-/-</sup>* samples to assess such a potential function under physiological conditions. We only considered as significant an absolute expression fold change greater than 1.5 and we set the false discovery rate (FDR) threshold at 5%. By using these parameters, we observed between 64 and 588 protein-coding genes differentially expressed in the various tissues analyzed (S3 and S4 Datasets). We first compared all tissues that express *Hotair* in the wild type



**Fig 4. Differential expression between wild type and *Del(Hotair)*<sup>-/-</sup> dissected samples.** A-B) Differential gene expression analysis between wild type and *Del(Hotair)*<sup>-/-</sup> trunk tissues (T1, T2, T3). The absolute fold change is > 1.5 and FDR < 0.05. The different columns correspond to sample type and rows correspond to differentially expressed genes. (A) Heat map of centered and scaled gene expression levels (Z-score log<sub>2</sub> RPKM). Genes are color coded vertically, according to tissue and expression changes between genotypes. (B) Venn diagram showing the number of down-regulated (top) and up-regulated (bottom) genes. C-D) Differential gene expression analysis between wild type and *Del(Hotair)*<sup>-/-</sup> forelimbs (FL), hindlimbs (HL) and genital tubercle (GT). The absolute fold change is > 1.5 and FDR < 0.05. (C) Heat map of centered and scaled gene expression levels (Z-score log<sub>2</sub> RPKM). Genes are color coded vertically according to the tissue and orientation of expression between genotypes. (D) Venn diagram showing the number of down-regulated (top) and up-regulated (bottom) genes.

doi:10.1371/journal.pgen.1006232.g004

condition, i.e. the T2, T3, GT and HL samples, reasoning that potential differentially expressed *Hotair* targets should be identified in these contexts. However, we were not able to identify any common genes with altered expression between wild type and *Del(Hotair)*<sup>-/-</sup> samples (S3 Fig), suggesting that the *Hotair* deletion does not affect the same set of genes in all tissues analyzed.

We thus divided the differential expression analysis based on the global expression clustering results. First, we compared the trunk samples T1 (lacking *Hotair* expression), T2 and T3. We identified 62 down-regulated genes and 13 up-regulated genes between wild type and *Del(Hotair)*<sup>-/-</sup> samples, which are shared in all trunk sections (Fig 4A and 4B). Of note, we observed a common trend in gene expression differences in all trunk samples, even though only some of them passed the established thresholds (Fig 4A). Gene ontology (GO) analysis for either common down-regulated or common up-regulated genes showed no enrichment of functional terms and the majority of differentially expressed genes were down-regulated, which was unexpected given the previously proposed role of *Hotair* as a repressor [20] (Fig 4A and 4B). GO analysis of down-regulated genes in distinct *Del(Hotair)*<sup>-/-</sup> trunk tissues revealed significant enrichment (FDR < 10%) in functional terms related to organ development and multicellular organismal process for most tissues (S5 Dataset). Up-regulated genes in T3 were

enriched for functional terms related to metabolic and biosynthetic processes and a weak enrichment for neuron differentiation genes was observed for T2 (S5 Dataset). Differential expression analysis for FL, HL and GT showed no common genes with altered expression (Fig 4C and 4D). GO analyses for differentially expressed genes in these tissues showed enrichment for functional terms related to development (S5 Dataset).

We next asked whether *Polycomb* target genes were preferentially up or down-regulated upon *Hotair* deletion. We defined putative target genes using H3K27me3 ChIP-seq data from wild type tail tip fibroblasts [20], selecting genes with a minimum 5-fold enrichment between H3K27me3 ChIP and input DNA in the gene promoter region (Materials and Methods) and thus obtained 861 putative target genes (S6 Dataset). We analyzed their pattern of differential expression in the T3 trunk sample, which includes the fetal tail and thus likely has the cell type composition *in vivo* most related to tail fibroblasts. Out of the 485 putative *Pc* targets that were expressed in the T3 segment (RPKM >1 in at least one wild type or *Del(Hotair)<sup>-/-</sup>* sample), 60 genes were significantly differentially expressed (absolute fold change > 1.5 and FDR <10%), including 50 down-regulated and 10 up-regulated in the *Del(Hotair)<sup>-/-</sup>* samples (S4 Fig). This indicates that only 17% of all differentially expressed *Pc* targets were up-regulated in *Del(Hotair)<sup>-/-</sup>* samples, which is slightly lower than the proportion of up-regulated genes among non-targets (25% up-regulated genes out of 564 differentially expressed non-target genes, Chi-square test p-value 0.18). Thus, we could not detect any enrichment for up-regulation of putative *Pc* target genes when compared to all other expressed protein-coding genes. Therefore, under these physiological conditions, we could not find evidence supporting a role for *Hotair* in setting up, maintaining or re-enforcing the repression of this set of *Polycomb* target genes. While physiologically relevant, our analysis is however difficult to directly compare with the situation in tail fibroblasts, as *Polycomb* occupancy naturally depends on both the tissue-type and the developmental stage.

### Expression of imprinted genes in *Del(Hotair)<sup>-/-</sup>* embryonic tissues

Interestingly, a subset of imprinted genes including *H19* and *Meg3* was shown to be up-regulated upon deletion of *Hotair* in TTF [20]. We thus analyzed the expression status of known imprinted genes transcribed (RPKM>1) in at least one sample (S7 Dataset). To ensure maximum sensitivity, we lowered our FDR threshold to 10% while maintaining an absolute expression fold change greater than 1.5. With these parameters, we observed a total of 21 imprinted genes differentially expressed in our samples (S5 Fig, S7 Dataset). We found that 71% of differentially expressed imprinted genes were down-regulated, while only 29% were up-regulated. Notably, *H19* and *Meg3* were down-regulated in our samples, in contrast to what was observed in TTF.

In conclusion, these global transcriptome analyses comparing *Hotair* deletion mutant and wild type micro-dissected tissues revealed changes in gene expression upon deletion of the *Hotair* locus. Noteworthy, we observed numerous expression changes not only in tissues that normally express *Hotair* at detectable levels, i.e. the T2, T3, GT and HL samples, but also in tissues like the anterior trunk (T1) or the forelimb (FL), where *Hotair* lncRNAs were not detected. This suggests that such observed differences in gene expression cannot be explained by a mere direct effect of the *Hotair* RNA. Potential explanations to these unexpected observations are discussed below.

### *Hox* genes expression in wild type and *Del(Hotair)<sup>-/-</sup>* mutant mice

The original observation, which led *Hotair* to become the paradigm of lncRNAs acting in *trans*, was its capacity to regulate several genes members of the *HoxD* cluster by interacting

with components of the *Polycomb* repressive complex 2 (PRC2)[13]. In contrast, no effect was initially reported upon *Hoxc* genes expression levels [13,20], despite the fact that *Hotair* is encoded from within the *HoxC* locus in both humans and mice [13,19]. We re-assessed this issue by analyzing the expression of all *Hox* genes across our various tissue samples (Fig 5). The global expression patterns of all four *Hox* clusters expectedly corresponded to previously described expression patterns for such embryonic stage and body levels (e.g. [24,25]).

In order to detect even subtle effects of *Hotair* upon *Hox* gene regulation, we lowered our FDR threshold to 10% for differential expression analyses. Under these conditions, we detected significant down-regulation of some anterior *Hoxa* genes (*Hoxa3*, *Hoxa5* and *Hoxa6*) in the three trunk samples (Fig 5 and S6 Fig). Interestingly, these differences were present in all trunk samples, including in T1 where *Hotair* is not expressed. We also observed a slight up-regulation of *Hoxb9* in HL and GT (Fig 5 and S6 Fig). Notably, in some of the tissues analyzed, we detected significant expression changes for both *Hoxc11* and *Hoxc12*, i.e. the two genes in the *HoxC* cluster that flank the *Hotair* locus (see below).

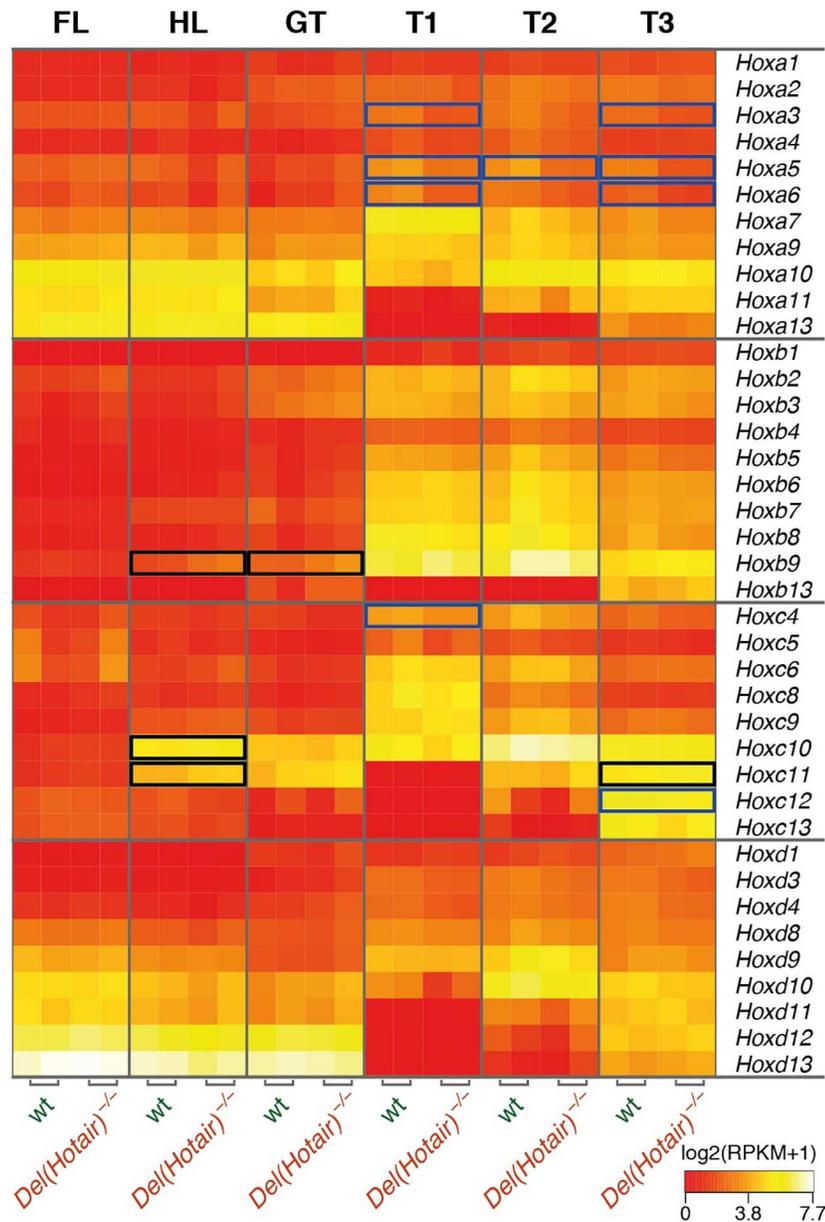
However, in contrast with previous reports using tail fibroblasts [13,20], we did not detect any significant change in the steady-state levels of *Hoxd* genes RNAs, in any of the analyzed tissues (Figs 5 and 6A). To clarify this contradictory observation, we re-analyzed the previously published RNA-seq data from both wild type and *Del(Hotair)<sup>-/-</sup>* TTF [20]. By implementing our analytical pipeline, we could not detect any significant difference in expression for any of the *Hoxd* genes (S7 Fig, S4 Dataset). Noteworthy, the expression levels of posterior *Hoxd* genes in this TTF dataset are either barely detectable or not detected at all, as for the *Hoxd12* gene, for example, suggesting that previous conclusions were raised based on particularly low transcript levels.

### Visualization of *Hoxd* genes expression in *Del(Hotair)<sup>-/-</sup>* embryos

The deletion of *Hotair* was claimed to alter both the expression levels and the spatial transcript distribution of the *Hoxd10* and *Hoxd11* genes in the trunk [20]. We performed whole mount *in situ* hybridization (WISH) on both wild type and *Del(Hotair)<sup>-/-</sup>* littermates to appreciate potential variations in the expression domains of these two genes. By using our well established protocol, we found that *Hoxd10* and *Hoxd11* transcripts showed wild type distributions in *Del(Hotair)<sup>-/-</sup>* mutant specimen (Fig 6B). To more precisely evaluate any potential difference in these expression domains between homozygous mutant and control littermates, we carried out double WISH for the *Hox* gene of interest in combination with a probe specific for the *MyoD* gene, which allowed for unambiguous somite visualization [26]. In both wild type and *Del(Hotair)<sup>-/-</sup>* embryos, *Hoxd10* was expressed in the future spine up to the level of somite 26, whereas *Hoxd11* was scored from somite level 29 and caudally, as previously reported [27]. Neither *Hoxd10*, nor *Hoxd11* showed any detectable increase in the intensity of the signal or in their spatial expression pattern, confirming the RNA-seq results (Fig 6A and 6B). Taken together, these observations suggest that *Hotair* has no effect on the regulation of *Hoxd* genes, at least in the developmental context and at the stage where the function of *Hox* genes is critical for morphological development.

### In-cis effect of *Hotair* deletion upon *Hoxc* genes transcription

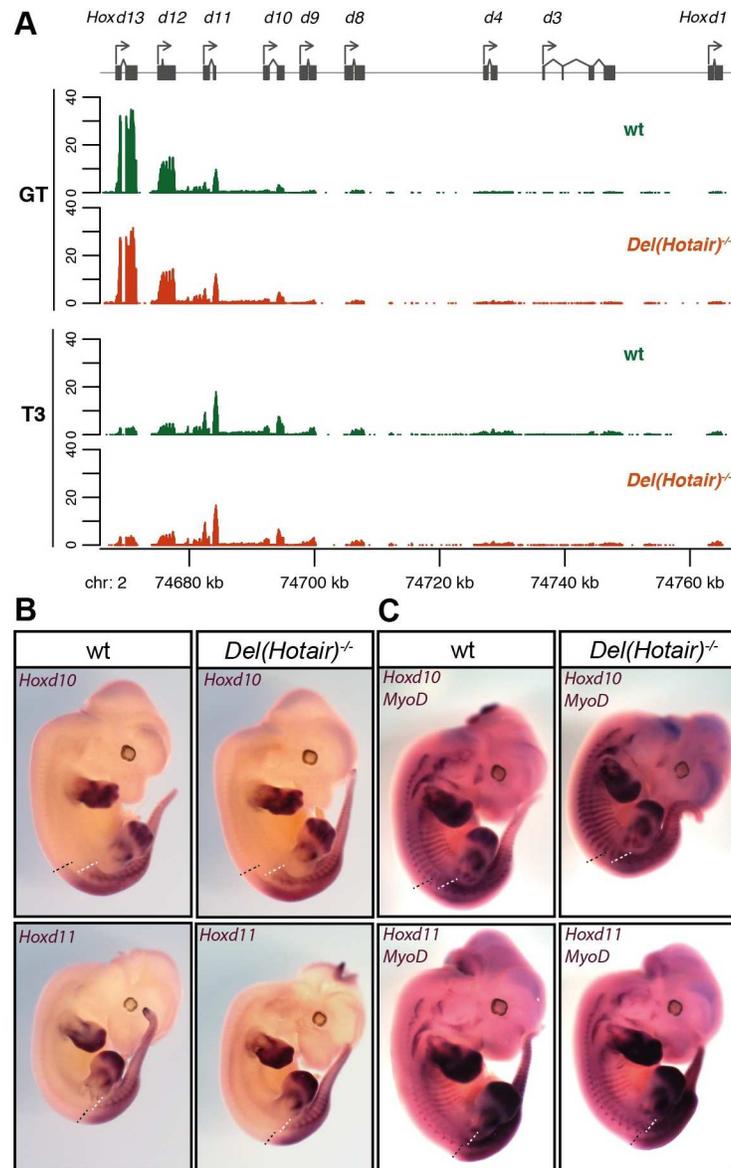
Unlike for *Hoxd* genes, our differential expression analyses between *Del(Hotair)<sup>-/-</sup>* and wild type samples revealed modest but significant changes for both *Hoxc11* and *Hoxc12*, the two genes neighboring the *Hotair* locus and thus flanking the deletion breakpoint (Fig 5). To further verify this new observation, we carefully analyzed the expression levels of both *Hoxc11* and *Hoxc12* in all tissue samples from where RNA-seq datasets had been obtained. As expected



**Fig 5. Expression of *Hox* genes in the various wild type and *Del(Hotair)*<sup>-/-</sup> embryonic tissues.** Heat map of log<sub>2</sub>-transformed RPKM expression levels for all *Hox* genes. The columns correspond to sample type (indicated on top) and the rows correspond to *Hox* genes (indicated on the right). The blue boxes point to down-regulated genes, whereas the black boxes indicate up-regulated genes (FDR < 10%, no minimal fold change threshold).

doi:10.1371/journal.pgen.1006232.g005

from their collinear transcription [28], *Hoxc12* transcripts were mainly detected in the most posterior T3 trunk sample. In this sample, a significant reduction in the level of *Hoxc12* RNAs was scored in *Del(Hotair)*<sup>-/-</sup> specimen (Fig 7A). On the other hand, *Hoxc11* transcripts were detected in the hindlimbs (HL), the genital tubercle (GT) and the T2 and T3 trunk samples. In these tissues, we observed an up-regulation of *Hoxc11* RNAs upon deletion of the *Hotair* locus, which was statistically significant for both HL and T3 (Fig 7B). In addition, a strong positive

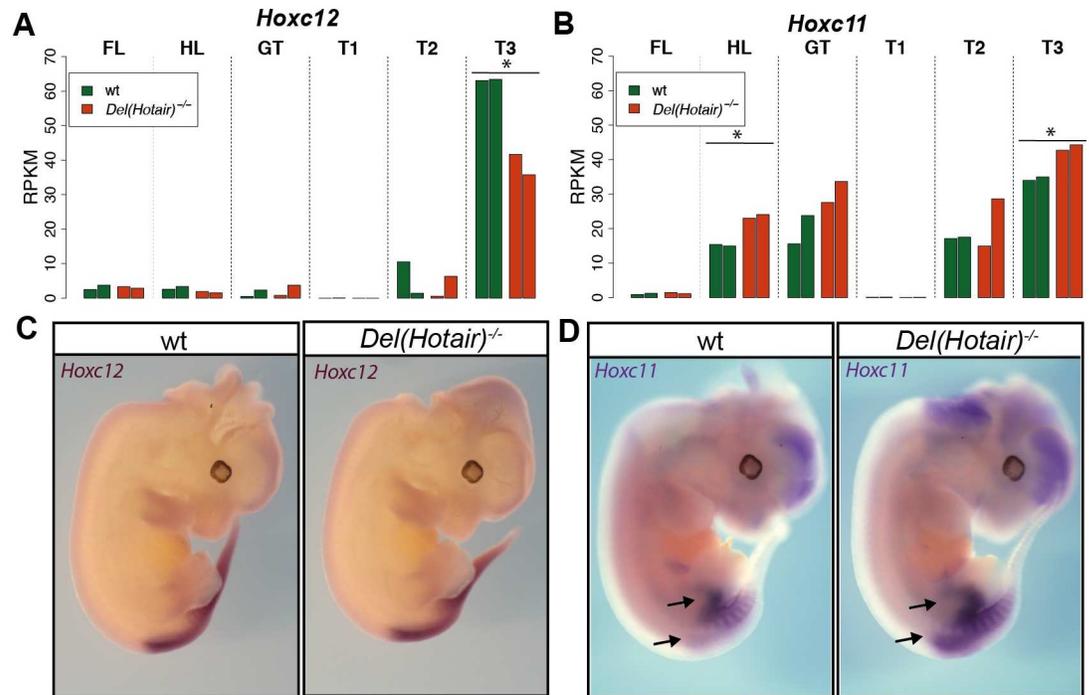


**Fig 6. The deletion of *Hotair* does not alter *Hoxd* genes expression in embryo.** (A) RNA-seq expression profiles of *Hoxd* genes in both the GT and T3 tissues of wild type (green) and *Del(Hotair)*<sup>-/-</sup> (orange) E12.5 embryos. The Y-axis represents the per-base read coverage, normalized by dividing by the total number of million mapped reads in the corresponding samples. The two biological replicates were pooled for this representation and only uniquely mapping reads were used. (B) WISH of *Hoxd10* and *Hoxd11* on E12.5 wild type (left) and *Del(Hotair)*<sup>-/-</sup> (right) embryos. The dashed lines indicate the rostral limits of the expression domains in the trunk, neural tube (black) and paraxial mesoderm (white). Adult vertebrae derive from the latter tissue. (C) Double WISH for the *MyoD* RNAs (for somite visualization) and either the *Hoxd10* (upper panel) or *Hoxd11* (lower panel) on E12.5 wild type (left) and *Del(Hotair)*<sup>-/-</sup> (right) embryos. There was no detectable difference in the anterior limit of expression for any *Hoxd* gene analyzed.

doi:10.1371/journal.pgen.1006232.g006

correlation between the expression levels of *Hotair* and *Hoxc11* was detected in wild type samples (S8 Fig). The correlation was weaker between *Hotair* and *Hoxc12* expression (S8 Fig).

We asked whether these changes in expression of *Hoxc* genes in some *Del(Hotair)*<sup>-/-</sup> samples were accompanied by alterations in their spatial expression patterns. We analyzed the expression of both *Hoxc12* and *Hoxc11* by WISH in *Del(Hotair)*<sup>-/-</sup> E12.5 embryos and wild type



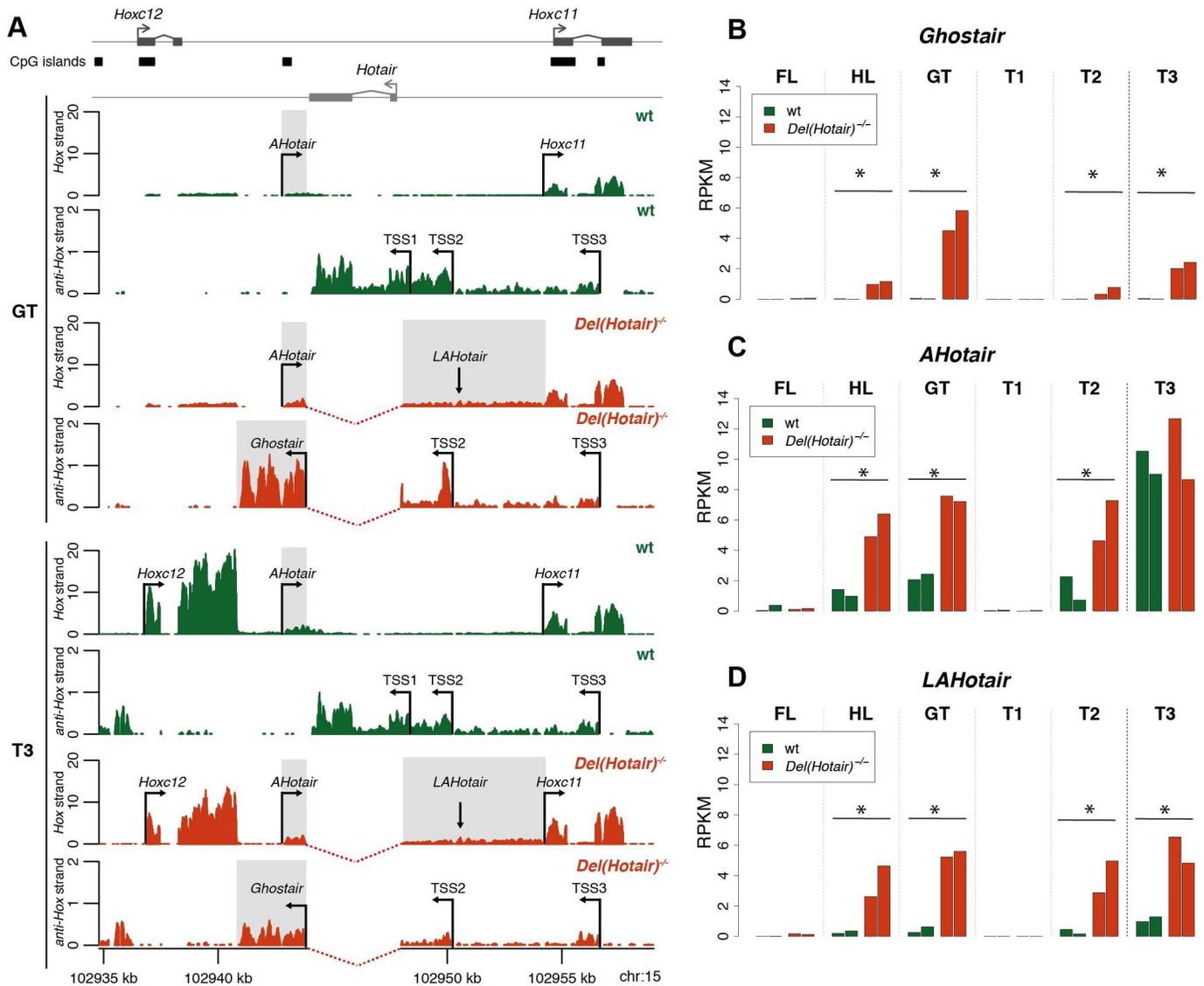
**Fig 7. The deletion of *Hotair* affects the expression of the neighboring *Hoxc11* and *Hoxc12* genes.** *Hoxc12* (A) and *Hoxc11* (B) expression (normalized RPKM values) in the various dissected tissue samples for wild type (green) and *Del(Hotair)<sup>-/-</sup>* (orange) E12.5 embryos. The asterisk\* indicates those samples where significant differences in transcript levels between genotypes were scored (FDR < 10%). (C) WISH using the *Hoxc12* probe in both wild type (left) and *Del(Hotair)<sup>-/-</sup>* (right) E12.5 embryos. The spatial expression of *Hoxc12* remains globally unchanged. (D) WISH of *Hoxc11* in wild type (left) and *Del(Hotair)<sup>-/-</sup>* (right) E12.5 embryos. The arrows indicate the slight anterior shift in the expression profile and the increase in signal intensity in the mutant genotype.

doi:10.1371/journal.pgen.1006232.g007

littermates. We did not observe any change for *Hoxc12* expression, neither in the transcript domain, nor in the intensity of RNA signal (Fig 7C). In contrast, *Del(Hotair)<sup>-/-</sup>* embryos showed a clear rostral expansion of the *Hoxc11* transcript domain in the trunk, as well as an increased signal intensity in both the hindlimb buds and the trunk, in agreement with our RNA-seq data (Fig 7D).

### The local impact of deleting the *Hotair* locus

To understand more precisely the reason why the deletion of the *Hotair* locus impacted the transcription of the flanking *Hoxc* genes, we analyzed in details the RNA-seq profiles of the region comprising *Hoxc12*, *Hotair* and *Hoxc11*. We first asked if all transcript isoforms derived from the *Hotair* locus were abrogated in the *Del(Hotair)<sup>-/-</sup>* allele and observed that the deleted region almost perfectly coincides with the annotated boundaries of the locus in the mouse. However, the inspection of the RNA-seq profiles in tissues that transcribe *Hotair* RNA revealed the existence of larger transcripts, extending over at least 2.4 kb upstream of the annotated promoter (Fig 8A). Although we cannot determine the precise location of *Hotair* transcription start site(s), the presence of continuous transcription upstream of the annotated gene boundaries indicates that at least one, and probably two of the possible *Hotair* promoters were not deleted. Indeed in *Del(Hotair)<sup>-/-</sup>* tissues, we detected transcripts initiating upstream of the annotated *Hotair* promoter, for instance within the *Hoxc11* intron, and spanning over the deleted region (Fig 8A).



**Fig 8. In-cis effects of the *Hotair* deletion on the local transcriptional activity.** A) RNA-seq expression profiles of the genomic region neighboring *Hotair* in both the developing genitalia (GT, four profiles on top) and the most posterior trunk tissue sample (T3, four profiles at the bottom) from either wild type (green) and *Del(Hotair)<sup>-/-</sup>* (orange) E12.5 embryos. In the wild type GT, only the *Hoxc11* gene is expressed along with *Hotair* on the opposite strand, which shows at least three putative start sites (arrows, TSS1 to TSS3). In the mutant GT, a long form of a new lncRNA (*AntiHotair*) now extends (grey box) on the *Hox* DNA strand, going over the deleted region up to the *Hoxc11* promoter. On the opposite DNA strand, the *Hotair* TSS2 and TSS3 are still functional and produce *Ghost of Hotair* (*Ghostair*), yet another new species of lncRNA, specific for the *Del(Hotair)<sup>-/-</sup>* allele (in orange) and absent from the control allele (in green). A similar situation is observed in the T3 trunk sample, except that *Hoxc12* and *AHotair* are also expressed there. In the native locus, *anti-Hotair* is produced and meets with the end of the *Hotair* transcript. In the deleted allele, *Ghostair* is produced by the remaining *Hotair* TSS and terminates close to the 3' end of the *Hoxc12* transcript (bottom two profiles). The gray boxes indicate the genomic regions used for the expression quantifications of *AHotair*, *LAHotair* and *Ghostair*. The Y-axis represents the per-base RNA-seq read coverage, normalized by dividing by the total number of million mapped reads in the corresponding samples. The two biological replicates were pooled for this representation and only uniquely mapping reads were used. B-D) Expression values (normalized RPKM) for *AHotair* (B), *LAHotair* (C) and *Ghost of Hotair* (D) in all tissue samples. Genotypes are color-coded with wild type in green and *Del(Hotair)<sup>-/-</sup>* in orange. The asterisk\* indicates those samples where significant differences in expression were scored between the two genotypes (FDR < 10%).

doi:10.1371/journal.pgen.1006232.g008

In contrast to its multiple start site(s), *Hotair* displayed a very sharp transcription termination site. In wild type tissues, transcription of *Hotair* terminated at the annotated site, with virtually no RNA-seq reads mapped downstream of this position (Fig 8A). However, in the *Del(Hotair)<sup>-/-</sup>* samples, we observed transcription downstream of the deleted locus terminating within 100bp of the *Hoxc12* termination site (Fig 8A). The presence of this extended transcript, which likely derives from one of the native *Hotair* promoters (as predicted with a *de novo* transcript assembly procedure, S9 Fig), likely resulted from the deletion of the wild type *Hotair* termination signals (Fig 8A). To quantify this gain of transcription, we counted RNA-seq reads mapping on the region between the annotated *Hotair* termination site and the *Hoxc12* termination site, on the *Hotair* strand. We referred to this transcript, which only appears upon deletion of *Hotair*, as *Ghost of Hotair* (*Ghostair*). We observed significant gains of *Ghostair* transcription in *Del(Hotair)<sup>-/-</sup>* samples, in all tissues that expressed *Hotair* in the wild type condition, i.e. the hindlimb buds, the genital bud and the two trunk samples T2 and T3 (Fig 8B).

### *Ghost of Hotair and Anti-Hotair*

Subsequent analyses of the RNA-seq profiles revealed an additional un-annotated promoter sequence, yet located on the *Hox* DNA strand. This promoter lies between *Hoxc12* and *Hoxc11* and overlaps with a CpG island (Fig 8A). In the wild type situation, it generates a relatively short, poorly abundant and un-spliced transcript, ca 1.8kb in size. The estimated termination site for this transcript was found within the region deleted in *Del(Hotair)*, close to the termination of *Hotair* itself on the other strand. Accordingly, we refer to this short transcript as *Anti-Hotair* (*AHotair*). In *Del(Hotair)<sup>-/-</sup>* samples, this CpG island promoter was still active, giving rise to a much longer *AHotair* transcript (Fig 8A, long *AHotair* or *LAHotair*), consistent with the deletion of its termination site. We did not observe any clear boundaries between this extended transcript and *Hoxc11*, suggesting that this *AHotair* RNA could leak onto the *Hoxc11* transcription unit. This was confirmed by a *de novo* transcript assembly procedure (see [Materials and Methods, S9 Fig](#)).

To quantify this gain of transcription from the *Hox* strand, we further defined two transcribed regions; The first one largely corresponded to the short *Anti-Hotair* transcript detected in wild type samples, starting at the CpG island promoter and ending at the boundary of the deleted region (Fig 8A). The second one, long *AHotair*, corresponded to the longer transcript observed in *Del(Hotair)<sup>-/-</sup>* samples, starting at the deleted region boundary and ending at the annotated *Hoxc11* transcription start site (Fig 8A). In agreement with our observations based on the RNA-seq profiles, we detected significant increases in expression for both *AHotair* and *LAHotair* in *Del(Hotair)<sup>-/-</sup>* mutant tissues (Fig 8C and 8D). Therefore, the deletion of the two annotated exons of *Hotair* [20] had a previously ignored important impact in *cis* by generating two new transcripts, which may potentially interact with the transcription of neighboring *Hoxc* genes.

## Discussion

### Homeotic *versus* Homeopathic phenotypes

In this study, we have re-investigated the phenotypic and molecular effects of deleting the *Hotair* lncRNA on mouse development *in vivo*, as reported in [20]. In this previous study, three phenotypic differences between wild type and *Del(Hotair)<sup>-/-</sup>* mice were reported, namely wrist malformation, a posterior homeotic transformation from lumbar vertebra L6 to sacral vertebra S1 identity and a mild anterior homeotic transformation of the 4th caudal vertebra. We did not detect any wrist malformation, nor did we see any substantial homeotic phenotype in the lumbar region of mutant animals, thus contradicting two of the three reported

phenotypic effects of the *Hotair* deletion. In *Mus musculus*, the lumbo-sacral transition shows great variability between L5 and L6 depending on the inbred strain considered and the total number of pre-sacral vertebrae. In fact, this number not only varies between inbred strains but also within the same strain and can even be biased by the sex of the animal [29]. Therefore, this region must be considered with great care before concluding on the presence of a homeotic transformation.

We note that another study involving two distinct deletion alleles of *Hotair*—though of smaller extents—also failed to confirm these latter two phenotypic effects [21]. The lack of effect of *Hotair* deletion upon wrist morphology is consistent with the absence of any detectable *Hotair* transcripts in mouse embryonic forelimbs (Fig 1) also reported previously [19] and by [21] using a sensitive *lacZ* reporter transgene system. Regarding the reported L5 to L6 transition, it is noteworthy that in wild type animals, detectable expression of *Hotair* in the paraxial mesoderm, i.e. in the mesodermal tissue that will generate the vertebrae, barely reaches the level of the lumbo-sacral transition, a transition labeled by its neighboring *Hoxc11* gene [30,31]. This makes a more anterior (at the L5 to L6 transition) *Hotair* loss-of-function dependent gain of function phenotype due to *Hoxd* genes difficult to understand.

Our analyses did nevertheless reveal a subtle difference between wild type and *Del(Hotair)*<sup>-/-</sup> mice in the morphology of the post-sacral caudal vertebrae. Although this observation is in agreement with one of the previously described morphological alterations [20,21], we note that the penetrance of the mutant phenotype is much lower than the 100% reported by [20]. Also, such an anterior transformation should reflect a loss of function rather than the effect of de-repressed *Hox* genes [1,3], as already scored in some instances, for example when abrogating the function of the nearby located *Hoxc13* gene [32]. Moreover, we also observed variations in these vertebral morphologies amongst wild type animals. A potential explanation for the observed difference in phenotypic penetrance may reside in the genetic background of the animals. In this work, we used a mixed CBAxBL/6 strain, while previous studies used a BL/6 background. This relatively mild difference in genetic backgrounds may account, at least in part, for the discrepancy regarding the penetrance of this weak and physiologically poorly significant morphological transformation. Should this be the case, we would still have to conclude that the previously reported phenotypic effects of *Hotair* deletion are not only very mild but also inbred strain-specific, definitely arguing against a general role—even minor—of *Hotair* during mouse development. Accordingly, we would refer to these phenotypic alterations as homeopathic rather than homeotic [20].

### The effects of the *Hotair* deletion in *trans*

Using sensitive RNA-seq measurements, we showed that the expression of hundreds of genes changed significantly upon deletion of the *Hotair* locus *in vivo*. However, none of these changes in gene expression could be reconciled with the suggested role for *Hotair* in silencing gene expression *in vivo* [20]. In particular, the initial proposal that *Hotair* RNA acts in *trans* to repress the expression of posterior *Hoxd* genes and of a subset of imprinted genes *via* the recruitment of the PRC2 complex [13,20] was not supported by our results. Indeed we did not note any significant change either in the levels, or in the spatial distribution of *Hoxd* transcripts, in any of the tissues analyzed. Also, when a larger set of putative *Pc* target genes was considered, the same conclusion was reached (S4 Fig). Finally, the majority (71%) of the differentially expressed imprinted genes including the reported *Hotair* targets *H19* and *Meg3*, were down- rather than up-regulated in *Del(Hotair)*<sup>-/-</sup> mutant samples, again in contradiction with previous results.

Therefore, our results are at odds not only with the phenotypic outcome of the *Hotair* deletion, but also with its effects upon gene expression [20]. One potential explanation to these

serious discrepancies may be that the regulatory effect of *Hotair* is highly specific for tail tip or foreskin fibroblasts, which were previously used for functional investigations [13,20], whereas not at work *in vivo*, precisely in those embryonic tissues where *Hotair* is expressed at the highest levels. Indeed both *Hotair* and *Hoxd* genes transcripts are rather abundant in our tissue samples, while they are very weakly present in murine tail tip fibroblasts (Fig 1, S7 Fig). Unless *Hotair* would function more efficiently at low concentrations, we conclude that our *in vivo* system is better suited to reveal the role of *Hotair*, if any.

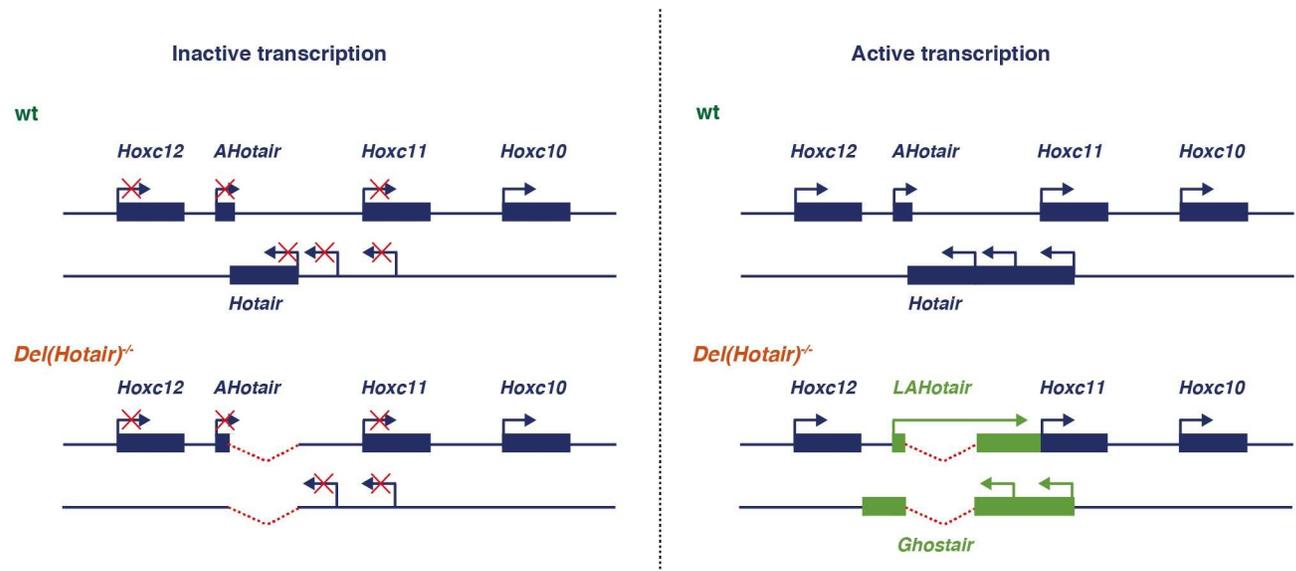
Another possibility is that the function of *Hotair* might not be exerted at the developmental stage analyzed (E12.5) but instead, at other time points. This explanation is nevertheless not compatible with the absence of phenotypic effects on skeletal morphology at P22, which should still be scored, should the deletion of *Hotair* deregulate target genes at other developmental stages. Also, the various genetic backgrounds may influence the penetrance of the phenotype (see above) and, by genotyping through the *Hotair* deleted locus, one may select for one particular haplotype associated with the mutant allele, which may result in some differential gene expression. Finally, it remains possible that a few hours difference in the developmental timing may lead to substantial relative variations in amounts of transcripts for many genes, in particular at an embryological stage where many important differentiation events occur.

In this context, it must be noted that the settings used for our transcriptome analyses overlap in sensitivity with the biological variations of the system itself, as seen for example with the variation in the level of *Hotair* in the GT replicate samples (Fig 1). Such differences can be due to intrinsic variations, yet most likely to slight variations in the micro-dissection plans or in the developmental stage of littermate embryos, or both. For example, a slight variation in the thickness of the piece in the trunk would elicit quantitative differences in *Hox* gene expression, whereas the depth of the piece (trunk) or the proximal level of the section (limbs, genitalia) may involve another presumptive tissue type, leading to large qualitative differences in transcripts. In fact, many of the strongest differentially expressed genes are clearly unrelated to those developmental processes involved in the potential morphological or molecular phenotypes under scrutiny (S10 Fig). This would also explain that differences are seen even in those samples where neither *Hotair*, nor *Hoxc11* are expressed. Accordingly, we do not interpret these results as reflecting changes in biological processes but, instead, as a sign of the sensitivity and intrinsic variations of our *in embryo* approach.

### The effects of the *Hotair* deletion in *cis*

When investigating the roles of lncRNAs by genetic approaches *in vivo*, it is often problematic to separate the lncRNA-dependent effects from those generated by the manipulation of the corresponding genomic locus [33]. *Hotair* is transcribed from within the *HoxC* cluster, a tightly packed and gene-dense locus, and its deletion was reported to have no consequence on the transcription of the neighboring *Hoxc* genes at the developmental stage and cell types examined [20]. Here again, our results *in embryo* contradict this view and showed that the expression levels of both *Hoxc11* and *Hoxc12* changed upon deleting the *Hotair* locus. We observed an extension in the spatial distribution of *Hoxc11* transcripts in both the trunk and the hindlimbs of *Del(Hotair)<sup>-/-</sup>* mutant specimens. Upon examination of the datasets of [20], we also found differences for *Hoxc10* and *Hoxc12* between wild type and *Del(Hotair)<sup>-/-</sup>* tail tip fibroblasts (S7 Fig). Therefore, the deletion of the *Hotair* locus had a significant impact in *cis* on *Hoxc* gene expression, in both *in vivo* and *in vitro* systems.

This was confirmed by the observation of *Ghost of Hotair* (*Ghostair*), a novel RNA produced by the anti-*Hox* strand in the deletion mutant allele. This transcript initiates at one of the alternative *Hotair* promoters, which was not included into the deletion, and terminates



**Fig 9. Schematic summarizing the data shown in Fig 8.** In a wild type situation and at a body level anterior to *Hoxc11* expression (from example in the upper lumbar region such as L3 or L4), *Hoxc10* is active whereas the whole posterior part of the *HoxC* cluster is repressed (top left). Once *Hoxc11* and *Hoxc12* become activated, in more posterior regions of the body, both the *Hotair* and the *AntiHotair* RNAs are produced, from the anti-*Hox* and *Hox* DNA strands, respectively (top, right). Upon deletion of the *Hotair* locus, the posterior *HoxC* cluster remains closed for transcription in the anterior parts of the body (bottom, left). In contrast, the activation of the *Hoxc11* region (bottom right) triggers the transcription of *Ghostair* on the opposite DNA strand, which meets the 3' end of the *Hoxc12* transcription unit, perhaps causing the light decrease in *Hoxc12* mRNAs. On the *Hox* strand, the anti-*Hotair* RNA can now cross over the deleted region and contribute to the transcription of *Hoxc11*, perhaps inducing the observed light gain of expression of the latter gene.

doi:10.1371/journal.pgen.1006232.g009

close to the 3' end of the *Hoxc12* transcript on the opposite strand. Our analysis also revealed the existence of *AntiHotair*, a previously un-annotated transcript on the *Hox* strand, derived from a CpG island promoter located close to the 3' end of *Hotair*. While in the wild type situation this transcript remains relatively short and ends within the region targeted by the *Hotair* deletion, a longer *AntiHotair* transcript was produced in *Del(Hotair)<sup>-/-</sup>* samples with no clear separation with *Hoxc11*. As a consequence, this transcript could leak onto *Hoxc11*, acting as an alternative 5' un-translated region, which gives a mechanistic basis for the light gain of *Hoxc11* expression in *Del(Hotair)<sup>-/-</sup>* tissues (Fig 9). Such in-cis effects on the local transcription landscape by deleting transcription termination signals on both strands are likely independent from any possible *Hotair* function.

### Hotair back into context

The *Hotair* lncRNA is transcribed from within the *HoxC* gene cluster [13], i.e. one of the most gene-dense and GC-rich regions of mammalian genomes [34]. Due to the particular regulatory strategy at work on the four *Hox* gene clusters [8], any endogenous or exogenous promoter present within such a gene cluster will be transcribed at the place and time where the neighboring *Hox* genes will be activated. The transcription of *Hotair* is no exception to this rule, for transcripts are found posteriorly, roughly matching the expression domains of *Hoxc11* or *Hoxc12*. While it is indeed possible that *Hotair* exerts a genuine function during development, for example by micro-tuning the transcription of *Hoxc* genes in cis, the question as to whether or not this RNA could be a mere by-product of the complex regulation occurring in the gene cluster remains open, in our opinion.

In any case, a potential mis-regulation of *Hoxc* genes should be carefully considered when investigating *Hotair* functions. This is desirable not only when studying developmental phenotypes [18,19,35], where it may represent a confounding factor due to the known roles of *Hoxc* genes there, but also when studying the roles of *Hotair* in other biological processes including human diseases. For instance, it was reported that *Hotair* is overexpressed in breast cancer and that this RNA regulates metastasis by reprogramming chromatin via *Polycomb* complexes [36]. Our analysis of expression data obtained from a cohort of cancer samples [37] revealed a strong positive correlation between *Hotair* and *Hoxc11* expression (S8 Fig), also observed in our mouse wild type samples (S8 Fig). Therefore, while *Hotair* may indeed be involved in a variety of cancer conditions, it is likely that its over-expression in cancer cells is accompanied by *Hoxc11* over-expression, which may again confound the observed phenotypes.

## Conclusion

Thus far, four different alleles have been studied, which partially or entirely removed the *Hotair* lncRNA and no consensus has been found regarding a potential function of this RNA during mouse development [19–21]. In our hands, *Hotair* has no function during mouse development, *a fortiori* when the regulation of *Hoxd* genes in *trans* is concerned. The deletion of the locus engineered by [20] induces modifications in the transcription of some *Hoxc* genes, through complex re-allocations of promoter and termination sites leading to novel RNA species. This mis-regulation of *Hoxc* gene transcription may have a slight effect upon some vertebral morphologies, yet this impact—if any—would be poorly penetrant and inbred strain-specific, i.e. of little interest for our understanding of developmental processes at large. Yet another allele would be necessary to solve these discrepancies, whereby the CRISPR-cas9 technology would help abrogate the *Hotair* transcription without substantially modifying the *in-cis* environment. At this point however, we do not see the urgency of increasing the number of mutant alleles at this locus, as confounds due to genetic background differences may always blur the resolution of such subtle effects.

## Materials and Methods

### Mouse strains

The *Del(Hotair)* mouse strain was described in [20] and kindly provided by Dr. H. Chang. Heterozygous mice were crossed back onto a mixed CBAx57/B6 background (Charles River). Wild type, heterozygous and homozygous mutant embryos were obtained by inter-crossing heterozygous mice. Genotyping was performed by PCR analysis on individual yolk sac lysates using the following primers:

Wild type (F)–CCTTATTCTCCCGGAGCCTAGC

Wild type (R)–CTGCCTCTGGCTCCACTCC

*Del(Hotair)*<sup>-/-</sup> (F)–CCTTGCCAACGTGTGGCTTCC

*Del(Hotair)*<sup>-/-</sup> (R)–CCAAGTCTACCGCTACTACTGGC

### Ethics statement

Maintenance of, and experiments on animals were approved by the Geneva Canton ethical regulation authority (authorization GE/81/14 to D.D.) and performed according to Swiss law.

## Whole-mount *in situ* hybridization

Whole-mount *in situ* hybridizations (WISH) were performed according to standard protocols. Embryos were dissected in PBS and fixed from overnight in 4% paraformaldehyde (PFA), washed in PBS, dehydrated and stored in 100% methanol at  $-20^{\circ}\text{C}$ . Both *Del(Hotair)<sup>-/-</sup>* and control wild type E12.5 littermates embryos were processed in parallel to maintain identical conditions throughout the WISH procedure. DIG-labeled probes for *in situ* hybridizations were produced by *in vitro* transcription (Promega) and detection was carried out using an alkaline phosphatase conjugated anti-digoxigenin antibody (Roche). WISH probes templates were previously described in: *Hotair* [19]; *Hoxd10* and *Hoxd11* [27]; *Hoxc11* [35]; *Hoxc12* [34] and *MyoD* [26].

## Skeletal Preparation

Whole mount skeletal preparation of P22 animals was done with standard Alcian blue/Alizarin red staining protocols.

## RNA extraction and RNA-seq library preparation

Embryonic tissues were stored at  $-80^{\circ}\text{C}$  in RNAlater stabilization reagent (Ambion) before genotyping. After genotyping and embryo sorting, total RNA was extracted from tissues using QIAGEN RNeasy Plus Micro Kit after disruption and homogenization. RNA quality was assessed using an Agilent 2100 Bioanalyser. Only samples with high RNA integrity number were used. Sequencing libraries were prepared according to TruSeq Stranded mRNA Illumina protocol, with polyA selection. RNA-seq libraries were sequenced on an Illumina HiSeq 2500 sequencer, as single-end reads (read length 100 base pairs). We obtained between 36 and 54 millions of raw RNA-seq reads for each sample (S1 Table).

## RNA-seq processing and gene expression estimation

Raw RNA-seq reads were aligned on the mouse mm10 genome assembly using TopHat 2.0.9 [38]. Gene expression computations were performed using uniquely mapping reads extracted from TopHat alignments and genomic annotations from Ensembl release 82 [39]. We filtered the annotated transcript isoforms for protein-coding genes, keeping only transcripts annotated as 'protein-coding', thus discarding transcripts flagged as 'retained\_intron', 'nonsense-mediated decay' etc. For *Hox* genes, we manually inspected annotated transcripts and retained only the canonical isoform for each gene, discarding read-through transcripts and retained introns. For non-coding genes, all annotated isoforms were kept. We then constructed 'flattened' gene models by combining the exon coordinates from all retained isoforms and counted the number of unique reads that aligned on these exons. We discarded reads that aligned on two or more overlapping genes on the same strand, as well as reads containing more than 2 mismatches or small insertions or deletions. We computed RPKM (Read per Kilobase of Exon per Million mapped reads) expression levels for each gene based on the unique read counts. The total number of mapped reads was computed on the entire nuclear genome, discarding reads that mapped on the mitochondria. RPKM expression levels were then further normalized across samples with a median scaling procedure, using as a standard the 100 genes with the least expression rank variation across samples, found in the 25%-75% range of expression levels [40]. As a control, we also computed expression levels using all TopHat mapped reads and the multi-read and fragment bias correction procedures implemented in Cufflinks [41]. The same procedure was applied for previously published tail tip fibroblast RNA-seq samples [20]. The RNA-seq data presented in this previous publication were also strand-specific and

generated with a dUTP protocol that sequences the antisense mRNA strand like the TruSeq Stranded mRNA protocol.

## Statistical analyses and graphical representations

The principal component analysis (PCA) was performed using the `dudi.pca` function in the `ade4` package in R [42]. The input table for the PCA consisted of log<sub>2</sub>-transformed RPKM expression levels, for all protein-coding genes that had RPKM > 1 in at least one of our samples. The data was centered (meaning that the mean expression levels were brought to a value of 0 for each gene, removing between-gene variations in expression levels) but not scaled prior to the PCA analysis. Euclidean distances between samples were computed with the standard `dist` function in R and clustered using the hierarchical clustering method (`hclust`). All statistics and graphical representations were done in R.

## Differential expression analyses

We tested for differential gene expression using DESeq2 [43] in R. Specifically, we contrasted a generalized linear model that explains the variation in read counts for each gene, as a function of the genotype (wild type or *Del(Hotair)*<sup>-/-</sup>) with a null model that assumes no effect of the genotype. The analyses presented in the manuscript were performed with the likelihood ratio test (LRT); the Wald test was performed as a control and results are provided in the supplementary datasets. The tests were performed separately for each tissue. The p-values were corrected for multiple testing with the Benjamini-Hochberg approach, for all six tissues at the same time. The same procedure was applied for previously published tail tip fibroblast RNA-seq samples, which included wild type, heterozygous *Del(Hotair)*<sup>+/-</sup> and homozygous *Del(Hotair)*<sup>-/-</sup> samples. In this case, we performed three separate pairwise comparisons between the three genotypes.

## Gene ontology enrichments

We tested for gene ontology (GO) enrichment in the sets of differentially expressed genes using the GOrilla webserver [44]. Each enrichment analysis compared two lists of genes, the focal list containing differentially expressed protein-coding genes (up-regulated and down-regulated genes analyzed separately) and the background list containing all protein-coding genes expressed in the corresponding samples. To construct the background list, we computed the minimum number of reads observed for differentially expressed protein-coding genes, summed across all relevant samples and selected all genes that had equivalent or higher read counts.

## Prediction of *Polycomb* target genes

To obtain a list of putative *Polycomb* target genes, we analyzed chromatin immuno-precipitation followed by sequencing (ChIP-seq) data for H3K27me3 and corresponding input data, from wild type tail tip fibroblasts [20]. We mapped the ChIP-seq data on the mm10 mouse genome using Bowtie 2 [45]. We removed identical ChIP-seq reads to avoid biases stemming from PCR duplication and we kept unambiguously mapped reads with at most two mismatches. We computed the average H3K27me3 and input read coverage in the promoter region (defined as 2kb upstream the annotated transcription start site) for each Ensembl-annotated transcript. The same conclusions were reached when defining promoter regions as 4kb regions centered on the TSS (S6 Dataset). The read coverage was normalized by dividing by the total number of million mapped reads for each sample. We defined putative *Polycomb* targets as those genes for which the ratio between the H3K27me3 and input was at least 5, and

for which the absolute H3K27me3 normalized coverage was at least 0.1. We discarded genes that had satellite repeats in the promoter regions as we observed that these repeats are enriched in H3K27me3 marks (likely as an artifact).

### Imprinted genes

The list of known mouse imprinted genes was extracted from <http://www.geneimprint.com>.

### De novo transcript assembly at the *HoxC* locus

We used RNA-seq data from our *Del(Hotair)<sup>-/-</sup>* samples, combined across all six tissues, to predict transcript sequences for *Ghostair* and *AntiHotair*. To do this, we first split each RNA-seq reads into three segments and aligned them with Bowtie 2 on the DNA sequence delimited by *Hoxc12* and *Hoxc11*. We then extracted all RNA-seq reads that mapped at least partially onto this region and assembled transcripts *de novo* using Trinity [46] setting `SS_lib_type = R` since our data was strand-specific. We kept Trinity contigs with a minimum length of 1000bp and aligned them on the mouse chromosome 15 using BLAT [47]. We manually excluded small, repetitive BLAT hits. See also [S8 Dataset](#).

### Expression data in cancer samples

To study the correlation between *Hotair* expression and the expression of neighboring genes *Hoxc11* and *Hoxc12*, we analyzed gene expression levels (RPM = reads per total million mapped reads) for a cohort of cancer samples [37].

## Supporting Information

**S1 Fig. Correlations between biological replicates.** The scatterplots show the correlation of log<sub>2</sub>-transformed RPKM expression levels of all expressed protein coding genes between biological replicates (Pearson and Spearman correlation coefficients). FL, Forelimbs; HL, hindlimbs; GT, genital tubercle; T1, T2, T3; trunk samples corresponding to either the lumbosacral, the sacro-caudal region or the caudal region, respectively.  
(TIF)

**S2 Fig. Pairwise Euclidean distances between samples.** Hierarchical clustering and heat map of pairwise Euclidean distances between samples, computed on log<sub>2</sub>-transformed RPKM expression levels of all protein coding genes. The distances are color-coded, with blue representing small distances and yellow large distances.  
(TIF)

**S3 Fig. Numbers of differentially expressed genes, in common among tissues.** Venn diagrams of differential expression analysis results (fold change > 1.5 and FDR < 0.05) for all tissue samples that express *Hotair* in the wild type condition (T2, T3, GT and HL). (A) Venn diagram showing the down-regulated genes. (B) Venn diagram showing up-regulated genes.  
(TIF)

**S4 Fig. Differential expression for candidate *Polycomb* target genes in the T3 trunk sample.** (A) Volcano plot representing the log<sub>2</sub> fold change and the false discovery rate (log<sub>10</sub>-transformed) for candidate *Polycomb* target genes (see [Materials and Methods](#)) in T3. The direction of expression changes is color-coded, with red showing down-regulated genes and green up-regulated genes. Non-significant genes are in blue. (B) Barplot representing the percentage of up-regulated and down-regulated genes for candidate *Polycomb* targets (left) and other genes (right).  
(TIF)

**S5 Fig. Expression of known imprinted genes in wild type and *Del(Hotair)*<sup>-/-</sup> embryonic tissues.** Heat map of log<sub>2</sub>-transformed RPKM expression levels of all imprinted genes (extracted from <http://www.geneimprint.com>) expressed in our samples (RPKM > 1, S7 Dataset). The columns correspond to samples and the rows correspond to imprinted genes. Blue boxes indicate down-regulated genes whereas black boxes indicate up-regulated genes (fold change > 1.5 and FDR < 10%).

(TIF)

**S6 Fig. Expression of *Hoxa* and *Hoxb* genes in wild type and *Del(Hotair)*<sup>-/-</sup> dissected embryonic tissues.** (A) RNA-seq expression profiles of the *HoxA* genomic region in the trunk T1 (top), T2 (middle) and T3 (bottom) samples of either wild type (green) or *Del(Hotair)*<sup>-/-</sup> (orange) E12.5 embryos. Very subtle differences are scored (arrows), which are nevertheless considered as significant using our analytical parameters (see also Fig 5). The Y-axis represents the per-base RNA-seq read coverage, normalized by dividing by the total number of million mapped reads in the corresponding samples. The two biological replicates were pooled for this representation and only uniquely mapping reads were used. (B) RNA-seq expression profiles of the *HoxB* genomic region in the hindlimbs (HL, top) and genital tubercle (GT, bottom) of either wild type (green) or *Del(Hotair)*<sup>-/-</sup> (red) E12.5 embryos. There again, the difference observed for *Hoxb9* (Fig 5) is weak yet considered as significant in our conditions (FDR < 10%, no fold change threshold).

(TIF)

**S7 Fig. Expression of *Hox* genes in tail tip fibroblasts (TTF).** Bar plots showing the quantification of all *Hox* genes expression by RNA-seq (normalized RPKM values) in TTF. Datasets are from [20]. The FDR of the differential expression test (likelihood ratio test in DESeq2) between wild type and *Del(Hotair)*<sup>-/-</sup> samples is indicated in red above each gene.

(TIF)

**S8 Fig. Correlations between the expression levels of *Hotair* and of the neighboring *Hoxc* genes.** (A) Scatterplot of log<sub>2</sub>-transformed RPKM expression levels for *Hotair* and *Hoxc11* shows excellent correlation (R = 0.898). (B) Scatterplot of log<sub>2</sub>-transformed RPKM expression levels for *Hotair* and *Hoxc12* with a lower correlation coefficient (R = 0.592). The various tissues samples are represented by a color code and the genotypes are indicated by either a circle (wild type), or a triangle (*Del(Hotair)*<sup>-/-</sup>) (C) Scatterplot of log<sub>2</sub>-transformed RPM expression levels for *Hotair* and the *Hoxc11* gene in a cohort of 376 cancer samples [37], showing a high correlation coefficient (R = 0.92). (D) Scatterplot of log<sub>2</sub>-transformed RPM expression levels for *Hotair* and the *Hoxc12* gene in the same cohort as in C)[37], showing a lower correlation coefficient (R = 0.66). The Pearson correlation coefficients are shown above the plot.

(TIF)

**S9 Fig. *De novo* transcript assembly in the *HoxC* cluster for a pool of the various samples in *Del(Hotair)*<sup>-/-</sup> mutant specimens.** (A) Existing annotations for *Hoxc11*, *Hoxc12* and *Hotair*, as extracted from the Ensembl database. (B) Variation in the RNA-seq reads coverage at the vicinity of *Hotair*. The *Hox* and anti-*Hox* strands are depicted separately. The Y-axis represents the per-base RNA-seq read coverage, normalized by dividing by the total number of million mapped reads in the corresponding samples. All our *Del(Hotair)*<sup>-/-</sup> samples were pooled for this representation and only uniquely mapping reads were used. (C) Genomic coordinates of *de novo* assembled transcripts, as predicted by Trinity on the basis of *Del(Hotair)*<sup>-/-</sup> RNA-seq data and mapped on the genome with Blat (excluding repetitive hits). The different isoforms assigned to a single gene were combined for this representation. Note that transcripts may be fragmented, in particular at repeats and low complexity regions. We observe transcripts

crossing the deleted region on both the *Hox* strand (Trinity identifier *TR5603|c0\_g3*) and the anti-*Hox* strand (Trinity identifier *TR5645|c0\_g1*). (D) Positions of repeated elements in the vicinity of *Hotair*.

(TIF)

**S10 Fig. Top differentially expressed genes in embryonic tissue samples.** The volcano plots show the log<sub>2</sub> fold change and the false discovery rate (log<sub>10</sub>-transformed) for differentially expressed genes (fold change > 1.5 and FDR < 0.05) in the various tissue samples. The top differentially expressed genes are highlighted in red (fold change > 4 and FDR < 0.00001). Dashed lines mark absolute fold changes of 2 and 4. FL, Forelimbs; HL, hindlimbs; GT, genital tubercle; T1, T2, T3; trunk samples corresponding to either the lumbo-sacral, the sacro-caudal region or the caudal region, respectively.

(TIF)

**S1 Dataset. Gene expression data for wild type and *Del(Hotair)*<sup>-/-</sup> embryonic mouse tissue samples.** Data are from forelimb buds (FL), hindlimb buds (HL), genital tubercle (GT) and the trunk T1, T2 and T3 samples. Numbers of uniquely mapped reads and RPKM/FPKM values are provided in separate files.

(TGZ)

**S2 Dataset. Gene expression for either wild type, *Del(Hotair)*<sup>+/-</sup> or *Del(Hotair)*<sup>-/-</sup> tail tip fibroblasts (TTF), after re-analysis of the RNA-seq datasets from Li and colleagues [20].** Numbers of uniquely mapped reads and RPKM/FPKM values are provided in separate files.

(TGZ)

**S3 Dataset. Differential expression analyses comparing wild type and *Del(Hotair)*<sup>-/-</sup> mouse tissue samples.** Data are from the FL, HL, GT, T1, T2 and T3 and are provided in separate files.

(TGZ)

**S4 Dataset. Differential expression analyses comparing wild type, *Del(Hotair)*<sup>+/-</sup> and *Del(Hotair)*<sup>-/-</sup> tail tip fibroblasts (TTF), after our re-analysis of the RNA-seq data from [20].**

(TGZ)

**S5 Dataset. Gene ontology (GO) enrichment analysis, contrasting genes that are differentially expressed between wild type and *Del(Hotair)*<sup>-/-</sup> mouse tissues from a background set of genes expressed at comparable levels in the same tissues.** The enrichment analysis was done for the GO category 'biological process', using the GORILLA webserver [44].

(TGZ)

**S6 Dataset. Analysis of putative *Polycomb* target genes, predicted on the basis of H3K27me3 ChIPSeq data from Li and colleagues [20].** The table provides the H3K27me3 and input read coverage for the predicted *Polycomb* targets, as well as differential expression results for these genes in the T3 trunk segment.

(TGZ)

**S7 Dataset. Results of the differential expression analyses for imprinted genes.** The list of imprinted genes was extracted from [www.geneimprint.com](http://www.geneimprint.com).

(TGZ)

**S8 Dataset. Results of the *de novo* transcript assembly procedure, which we used to confirm that *Ghostair* and *AntiHotair* cross the deleted region.**

(TGZ)

**S1 Table. Basic statistics for the RNA-seq dataset, including the number of raw and mapped RNA-seq reads and the number of detected genes for each sample.**  
(XLSX)

## Acknowledgments

We thank H. Chang for providing the *Hotair* mutant stock. We also thank M. Ros as well as L. Deslile and other lab members for their expertise, comments and materials, T.-H. Nguyen Huyn for mouse breeding, A. Awgulewitsch for WISH probes and J. Zakany for comments on the manuscript. We thank the Geneva Genomics Platform (University of Geneva) for sequencing.

## Author Contributions

**Conceived and designed the experiments:** ARA AN DD.

**Performed the experiments:** ARA AN EJ BM.

**Analyzed the data:** ARA AN DD.

**Wrote the paper:** ARA AN DD.

## References

- Lewis EB (1978) A gene complex controlling segmentation in *Drosophila*. *Nature* 276: 565–570. PMID: [103000](#)
- Krumlauf R (1994) Hox genes in vertebrate development. *Cell* 78: 191–201. PMID: [7913880](#)
- Duboule D, Morata G (1994) Colinearity and functional hierarchy among genes of the homeotic complexes. *Trends Genet* 10: 358–364. PMID: [7985240](#)
- Wellik DM (2009) Hox Genes and Vertebrate Axial Pattern. *Hox Genes* 88: 257–+.
- Holland PW, Garcia-Fernandez J (1996) Hox genes and chordate evolution. *Dev Biol* 173: 382–395. doi: [10.1006/dbio.1996.0034](#) PMID: [8605999](#)
- Duboule D (2007) The rise and fall of Hox gene clusters. *Development* 134: 2549–2560. doi: [10.1242/dev.001065](#) PMID: [17553908](#)
- Deschamps J, van Nes J (2005) Developmental regulation of the Hox genes during axial morphogenesis in the mouse. *Development* 132: 2931–2942. doi: [10.1242/dev.01897](#) PMID: [15944185](#)
- Kmita M, Duboule D (2003) Organizing axes in time and space; 25 years of colinear tinkering. *Science* 301: 331–333. doi: [10.1126/science.1085753](#) PMID: [12869751](#)
- Berlivet S, Paquette D, Dumouchel A, Langlais D, Dostie J, et al. (2013) Clustering of tissue-specific sub-TADs accompanies the regulation of HoxA genes in developing limbs. *PLoS Genet* 9: e1004018. doi: [10.1371/journal.pgen.1004018](#) PMID: [24385922](#)
- Soshnikova N, Duboule D (2009) Epigenetic temporal control of mouse Hox genes in vivo. *Science* 324: 1320–1323. doi: [10.1126/science.1171468](#) PMID: [19498168](#)
- Noordermeer D, Leleu M, Splinter E, Rougemont J, De Laat W, et al. (2011) The dynamic architecture of Hox gene clusters. *Science* 334: 222–225. doi: [10.1126/science.1207194](#) PMID: [21998387](#)
- Wang KC, Yang YW, Liu B, Sanyal A, Corces-Zimmerman R, et al. (2011) A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* 472: 120–124. doi: [10.1038/nature09819](#) PMID: [21423168](#)
- Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, et al. (2007) Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* 129: 1311–1323. doi: [10.1016/j.cell.2007.05.022](#) PMID: [17604720](#)
- Sauvageau M, Goff LA, Lodato S, Bonev B, Groff AF, et al. (2013) Multiple knockout mouse models reveal lincRNAs are required for life and brain development. *Elife* 2: e01749. doi: [10.7554/eLife.01749](#) PMID: [24381249](#)

15. Mainguy G, Koster J, Woltering J, Jansen H, Durston A (2007) Extensive polycistronism and antisense transcription in the mammalian Hox clusters. *PLoS One* 2: e356. doi: [10.1371/journal.pone.0000356](https://doi.org/10.1371/journal.pone.0000356) PMID: [17406680](https://pubmed.ncbi.nlm.nih.gov/17406680/)
16. Lee JT (2012) Epigenetic regulation by long noncoding RNAs. *Science* 338: 1435–1439. doi: [10.1126/science.1231776](https://doi.org/10.1126/science.1231776) PMID: [23239728](https://pubmed.ncbi.nlm.nih.gov/23239728/)
17. Tsai MC, Manor O, Wan Y, Mosammamaparast N, Wang JK, et al. (2010) Long noncoding RNA as modular scaffold of histone modification complexes. *Science* 329: 689–693. doi: [10.1126/science.1192002](https://doi.org/10.1126/science.1192002) PMID: [20616235](https://pubmed.ncbi.nlm.nih.gov/20616235/)
18. Suemori H, Noguchi S (2000) Hox C cluster genes are dispensable for overall body plan of mouse embryonic development. *Dev Biol* 220: 333–342. doi: [10.1006/dbio.2000.9651](https://doi.org/10.1006/dbio.2000.9651) PMID: [10753520](https://pubmed.ncbi.nlm.nih.gov/10753520/)
19. Schorderet P, Duboule D (2011) Structural and functional differences in the long non-coding RNA hotair in mouse and human. *PLoS Genet* 7: e1002071. doi: [10.1371/journal.pgen.1002071](https://doi.org/10.1371/journal.pgen.1002071) PMID: [21637793](https://pubmed.ncbi.nlm.nih.gov/21637793/)
20. Li L, Liu B, Wapinski OL, Tsai MC, Qu K, et al. (2013) Targeted disruption of Hotair leads to homeotic transformation and gene derepression. *Cell Rep* 5: 3–12. doi: [10.1016/j.celrep.2013.09.003](https://doi.org/10.1016/j.celrep.2013.09.003) PMID: [24075995](https://pubmed.ncbi.nlm.nih.gov/24075995/)
21. Lai KM, Gong G, Atanasio A, Rojas J, Quispe J, et al. (2015) Diverse Phenotypes and Specific Transcription Patterns in Twenty Mouse Lines with Ablated LincRNAs. *PLoS One* 10: e0125522. doi: [10.1371/journal.pone.0125522](https://doi.org/10.1371/journal.pone.0125522) PMID: [25909911](https://pubmed.ncbi.nlm.nih.gov/25909911/)
22. Tschopp P, Duboule D (2011) A genetic approach to the transcriptional regulation of hox gene clusters. *Annu Rev Genet* 45: 145–166. doi: [10.1146/annurev-genet-102209-163429](https://doi.org/10.1146/annurev-genet-102209-163429) PMID: [22060042](https://pubmed.ncbi.nlm.nih.gov/22060042/)
23. Tschopp P, Duboule D (2014) The Genetics of Murine Hox Loci: TAMERE, STRING, and PANTHERE to Engineer Chromosome Variants. *Methods Mol Biol* 1196: 89–102. doi: [10.1007/978-1-4939-1242-1\\_6](https://doi.org/10.1007/978-1-4939-1242-1_6) PMID: [25151159](https://pubmed.ncbi.nlm.nih.gov/25151159/)
24. Wellik DM (2007) Hox patterning of the vertebrate axial skeleton. *Dev Dyn* 236: 2454–2463. doi: [10.1002/dvdy.21286](https://doi.org/10.1002/dvdy.21286) PMID: [17685480](https://pubmed.ncbi.nlm.nih.gov/17685480/)
25. Zakany J, Duboule D (2007) The role of Hox genes during vertebrate limb development. *Current Opinion in Genetics & Development* 17: 359–366.
26. Hebrok M, Fuchtbauer A, Fuchtbauer EM (1997) Repression of muscle-specific gene activation by the murine Twist protein. *Exp Cell Res* 232: 295–303. doi: [10.1006/excr.1997.3541](https://doi.org/10.1006/excr.1997.3541) PMID: [9168805](https://pubmed.ncbi.nlm.nih.gov/9168805/)
27. Gerard M, Chen JY, Gronemeyer H, Chambon P, Duboule D, et al. (1996) In vivo targeted mutagenesis of a regulatory element required for positioning the Hoxd-11 and Hoxd-10 expression boundaries. *Genes Dev* 10: 2326–2334. PMID: [8824591](https://pubmed.ncbi.nlm.nih.gov/8824591/)
28. Izpisua-Belmonte JC, Falkenstein H, Dolle P, Renucci A, Duboule D (1991) Murine genes related to the *Drosophila* AbdB homeotic genes are sequentially expressed during development of the posterior part of the body. *EMBO J* 10: 2279–2289. PMID: [1676674](https://pubmed.ncbi.nlm.nih.gov/1676674/)
29. Green EL (1941) Genetic and Non-Genetic Factors Which Influence the Type of the Skeleton in an Inbred Strain of Mice. *Genetics* 26: 192–222. PMID: [17247002](https://pubmed.ncbi.nlm.nih.gov/17247002/)
30. Wellik DM, Capecchi MR (2003) Hox10 and Hox11 genes are required to globally pattern the mammalian skeleton. *Science* 301: 363–367. doi: [10.1126/science.1085672](https://doi.org/10.1126/science.1085672) PMID: [12869760](https://pubmed.ncbi.nlm.nih.gov/12869760/)
31. Hostikka SL, Capecchi MR (1998) The mouse Hoxc11 gene: genomic structure and expression pattern. *Mech Dev* 70: 133–145. PMID: [9510030](https://pubmed.ncbi.nlm.nih.gov/9510030/)
32. Godwin AR, Capecchi MR (1998) Hoxc13 mutant mice lack external hair. *Genes Dev* 12: 11–20. PMID: [9420327](https://pubmed.ncbi.nlm.nih.gov/9420327/)
33. Bassett AR, Akhtar A, Barlow DP, Bird AP, Brockdorff N, et al. (2014) Considerations when investigating lncRNA function in vivo. *Elife* 3: e03058. doi: [10.7554/eLife.03058](https://doi.org/10.7554/eLife.03058) PMID: [25124674](https://pubmed.ncbi.nlm.nih.gov/25124674/)
34. Peterson RL, Papenbrock T, Davda MM, Awgulewitsch A (1994) The murine Hoxc cluster contains five neighboring AbdB-related Hox genes that show unique spatially coordinated expression in posterior embryonic subregions. *Mech Dev* 47: 253–260. PMID: [7848872](https://pubmed.ncbi.nlm.nih.gov/7848872/)
35. Papenbrock T, Visconti RP, Awgulewitsch A (2000) Loss of fibula in mice overexpressing Hoxc11. *Mech Dev* 92: 113–123. PMID: [10727851](https://pubmed.ncbi.nlm.nih.gov/10727851/)
36. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, et al. (2010) Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* 464: 1071–1076. doi: [10.1038/nature08975](https://doi.org/10.1038/nature08975) PMID: [20393566](https://pubmed.ncbi.nlm.nih.gov/20393566/)
37. Balbin OA, Malik R, Dhanasekaran SM, Prensner JR, Cao X, et al. (2015) The landscape of antisense gene expression in human cancers. *Genome Res* 25: 1068–1079. doi: [10.1101/gr.180596.114](https://doi.org/10.1101/gr.180596.114) PMID: [26063736](https://pubmed.ncbi.nlm.nih.gov/26063736/)

38. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, et al. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14: R36. doi: [10.1186/gb-2013-14-4-r36](https://doi.org/10.1186/gb-2013-14-4-r36) PMID: [23618408](https://pubmed.ncbi.nlm.nih.gov/23618408/)
39. Yates A, Akanni W, Amode MR, Barrell D, Billis K, et al. (2016) Ensembl 2016. *Nucleic Acids Res* 44: D710–716. doi: [10.1093/nar/gkv1157](https://doi.org/10.1093/nar/gkv1157) PMID: [26687719](https://pubmed.ncbi.nlm.nih.gov/26687719/)
40. Brawand D, Soumillon M, Necsulea A, Julien P, Csardi G, et al. (2011) The evolution of gene expression levels in mammalian organs. *Nature* 478: 343–348. doi: [10.1038/nature10532](https://doi.org/10.1038/nature10532) PMID: [22012392](https://pubmed.ncbi.nlm.nih.gov/22012392/)
41. Roberts A, Trapnell C, Donaghey J, Rinn JL, Pachter L (2011) Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome Biol* 12: R22. doi: [10.1186/gb-2011-12-3-r22](https://doi.org/10.1186/gb-2011-12-3-r22) PMID: [21410973](https://pubmed.ncbi.nlm.nih.gov/21410973/)
42. Dray S, Dufour AB (2007) The ade4 package: Implementing the duality diagram for ecologists. *Journal of Statistical Software* 22: 1–20.
43. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15: 550. doi: [10.1186/s13059-014-0550-8](https://doi.org/10.1186/s13059-014-0550-8) PMID: [25516281](https://pubmed.ncbi.nlm.nih.gov/25516281/)
44. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z (2009) GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* 10: 48. doi: [10.1186/1471-2105-10-48](https://doi.org/10.1186/1471-2105-10-48) PMID: [19192299](https://pubmed.ncbi.nlm.nih.gov/19192299/)
45. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9: 357–359. doi: [10.1038/nmeth.1923](https://doi.org/10.1038/nmeth.1923) PMID: [22388286](https://pubmed.ncbi.nlm.nih.gov/22388286/)
46. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, et al. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 29: 644–652. doi: [10.1038/nbt.1883](https://doi.org/10.1038/nbt.1883) PMID: [21572440](https://pubmed.ncbi.nlm.nih.gov/21572440/)
47. Kent WJ (2002) BLAT—the BLAST-like alignment tool. *Genome Res* 12: 656–664. doi: [10.1101/gr.229202](https://doi.org/10.1101/gr.229202) PMID: [11932250](https://pubmed.ncbi.nlm.nih.gov/11932250/)