Recent experiments raised doubt on the vertebrate Temporal Collinearity of Hox Genes. A biophysical model may reconcile the conflicting findings.

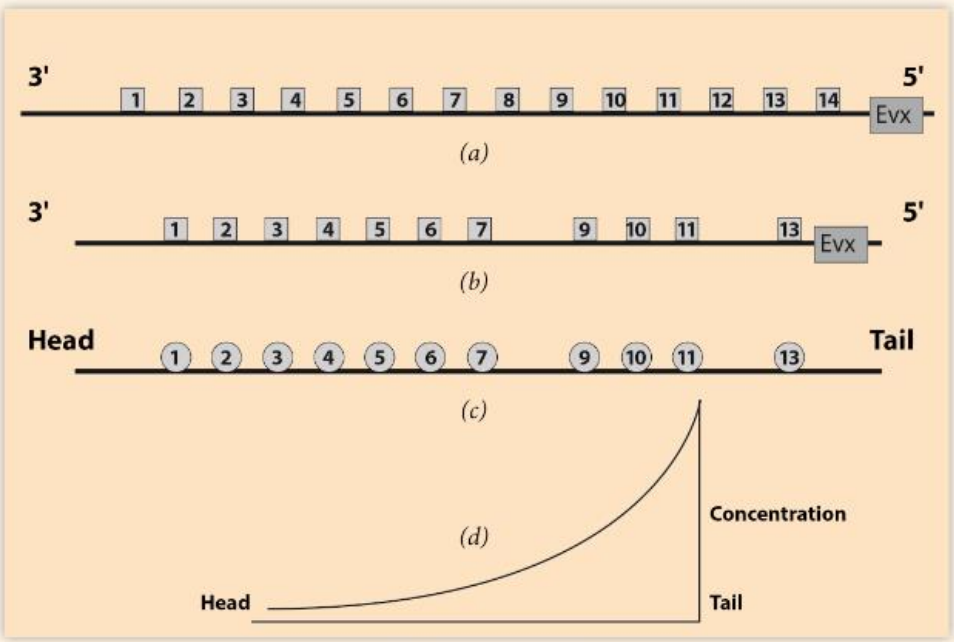
Spyros Papageorgiou
Institute of Biosciences and Applications, National Center for Scientific Research 'Demokritos' 153 10 Athens, Greece

Abstract

Hox gene collinearity (HGC) is a multiscalar property of many animal phyla particularly important during embryogenesis. It relates events occurring in Hox clusters inside the chromosome DNA and embryonic tissues. These two entities differ in size by more than four orders of magnitude. HGC is observed as spatial collinearity (SC) where the Hox genes are located in the order H1, H2, H3 … along the 3' to 5' direction of the DNA sequence. The corresponding embryonic tissues (E1, E2, E3, …) are activated along the Anterior – Posterior axis in the same order. Besides this collinearity a temporal collinearity (TC) has been also observed in many vertebrates. According to TC first is H1 expressed in E1, later is H2 in E2, followed by H3,… Lately doubt has been raised whether TC really exists. A biophysical model (BM) has been formulated and tested in the last twenty years. According to BM, physical forces are created which pull the Hox genes one after the other driving them to a transcription factory domain where they are transcribed. The existing experiments support this BM description. In the present work two equivalent realizations of BM are presented which explain the recent findings on TC as observed in the vertebrates.

1 Introduction

Hox Gene Collinearity (HGC) is a fundamental embryonic property coordinating development of vertebrates and many other animal phyla. It was first observed by E.B. Lewis in the Drosophila BX-C gene complex [1]. Lewis noticed that a class of Hox genes are located in clusters following an ordered sequence (Hox1, Hox2, Hox3 etc) on the chromosome (Fig.1). These genes are expressed in the same order in the embryo along the anterior/posterior (A/P) axis. This common order in the chromosome and the embryo is denoted spatial colinearity (SC).



[Figure 1]

Fig. 1: Hox genes in the Hox cluster (adapted from Papageorgiou S. [J-Multidisc.Sci. J.2020: 3(2), 151].. A) Hox gene ordering of a (theoretical) common ancestor. Gene Evx is located next to the 5' end of the Hox cluster. B) Ordering of the mouse HoxA cluster. HoxA8 and HoxA12 are missing. C) The corresponding ontogenetic units of the mouse along the A/P axis. D) The steady state monotonic concentration gradient of a morphogen. The peak is located at the tail region.

In the evolutionary process, a Hox cluster may appear in several homologues as a result of whole genome duplications (WGD) [5]. For instance, in vertebrates there are four homologue Hox clusters denoted HoxA, HoxB, HoxC, HoxD. These homologue clusters cooperate for the normal embryonic development. The ordered Hox genes of a cluster (Hox1, Hox2 ,..) constitute a paralogy group [4, 5]. For instance Pg1 can be traced in Hox clusters of different animal as a result of their similarity to the common ancestor Hox1. In vertebrates the paralogy group consists of 13 paralogue Hox genes Pg1, Pg2…,Pg13. In many cases, due to whole genome duplications, the genomes contain several copies of the Hox clusters. In a comprehensive study of Hox gene expressions in *Xenopus laevis*, M. Kondo et al. analyze the WGD and in particular the allotetraploidization which they estimated occurring 17 -18 Mya [5]. They designate the two subgenomes 'homologs' L and S and compare the homolog gene expressions within a Hox gene cluster. They conclude that many L, S homologs are not orderly correlated with the paralogue

order as would be expected according to TC. They conclude that 'the TC hypothesis must be revisited by comprehensive analysis of the developmental timing of transcriptional initiation of Hox genes…' [5] (See Section 3).

From a different point of view, the above doubt of TC validity is challenged by Durston et al. indicating that TC is verified in many vertebrates including cephalochordates [6, 7]. In particular, TC in cooperation with time-space transformations leads to the development of different body parts of the vertebrate embryo [7].

In view of the above ambiguity, it is worth interpreting the recent contradicting data applying a Biophysical Model (BM) which provides a unifying approach.

2 The Biophysical model for Hox Gene activation

The multiscale nature of HGC motivated the formulation of the following Biophysical Model (BM). The basic hypothesis is that pulling physical forces act on the Hox clusters. At the embryonic (macroscopic) scale the contribution on the forces is contained in a morphogen gradient along the Anterior-Posterior embryonic axis [8-10]. Before activation the Hox cluster is packed inside the chromatin territory (CT) in a compact unity. No forces are created at this stage. Gene activation starts when attractive forces are gradually created. Such forces emerge when polar molecules (positively charged) are allocated at the telomeric region of the cluster (Fig.2). (Detailed experimental evidence for these events is found in [9-11]). The forces gradually increase pulling the genes out of the CT (Fig.1). Tentatively, the nature of the forces is electric [8,9].

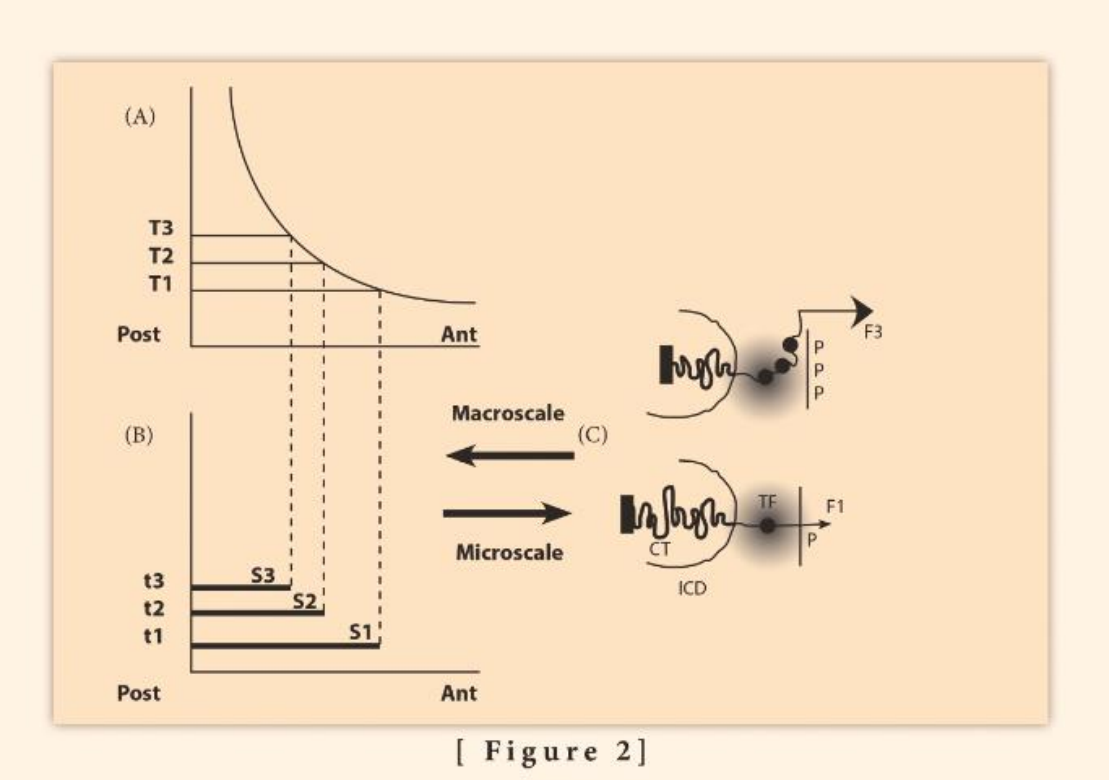


Fig.2: The macroscale morphogen gradient and the microscale Hox gene clustering in space and time (adapted from Papageorgiou S. [Biology 2017: 6, 32]). A) Concentration thresholds (T1,T2,T3) divide A/P axis in partially overlapping expression domains. B) The time sequence (t1, t2, t3) combined with thresholds sequence (T1, T2, T3) determine the Hox1, Hox2, Hox3 activation in space and time. S1, S2, S3 are the partially overlapping and nested expression domains of Hox1, Hox2, Hox3. C) (bottom) In an anterior cell of S1, a small force **F1** pulls Hox1 (black spot) out of CT toward the Interchromosome Domain (ICD) in the regime of the Transcription Factory (TF) (grey domain). Apposition of polar molecules P opposite the telomeric end of the Hox cluster (top). At a later stage in a more posterior location of S3, a stronger force **F3** pulls Hox1, Hox2, Hox3 out of CT in TF. Apposition of PPP molecules.

A simple heuristic form for the attractive forces **F** acting on Hox clusters is the following [10,11]

$$\mathbf{F} = \mathbf{P} \times \mathbf{N} \qquad (1)$$

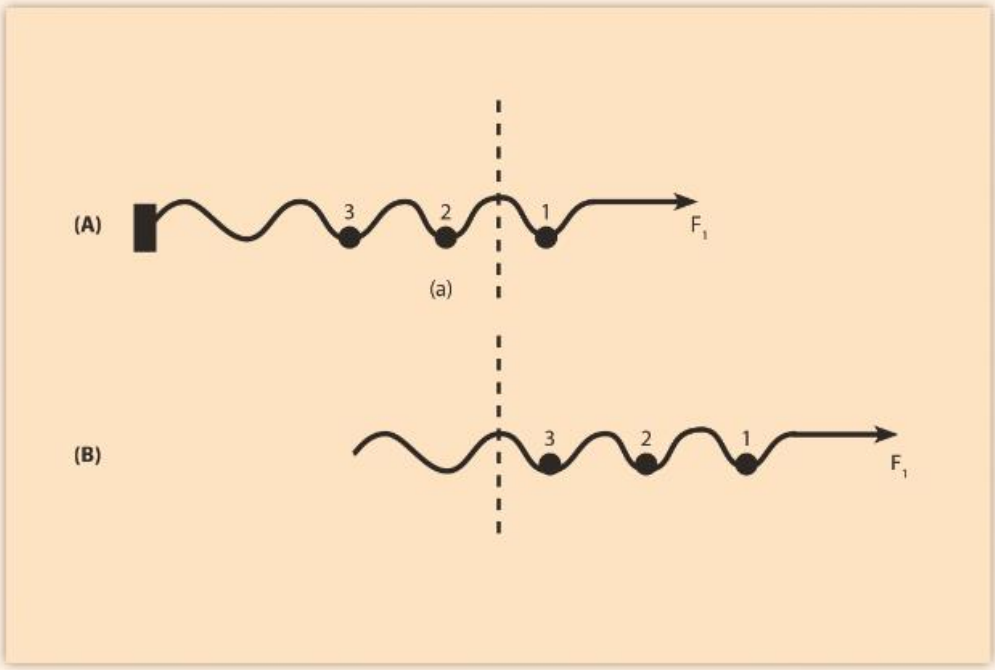
In eq. (1) factor **N** represents the contribution of the Hox cluster which is negatively charged in accordance with the overall negative charge of DNA. The factor **P** is graded (low anteriorily-high posteriorily) and represents the embryonic contribution to **F**. The force **F** is an electric quasi- Coulomb force and applies at the telomeric end of the Hox cluster (Fig. 2). BM and the attractive force **F** provide an interplay between the microscopic scale of the Hox cluster and the macroscopic embryonic scale (Fig. 2).

3 The expanding spring approximation of BM

With the development of novel technological methods like superresolution imaging of stochastic optical reconstruction microscopy (STORM) it has been made possible to measure the geometric modifications of hox clusters during Hox gene expressions [12-14]. It was thus found that the Hox clusters during gene activation are gradually elongated across the 3' to 5' axis of the cluster. This observation supports the hypothesis that the Hox clusters behave like an expanding elastic spring. BM predicts this behavior [15-17] (Fig.2). Following an increase of the morphogen gradient (from head to tail) the pulling force **F** increases and the number of extruded genes increases accordingly (Fig.2).

The mechanical properties of the expanding spring depend mainly on the fastening of the spring at the posterior end of the cluster and the 'stiffness' of the spring itself related to the local interactions of the constituent chromosomal configuration [18].

In the particular case of the mouse HoxD cluster, the fastening domain of the HoxD cluster is contained in the chromosome domain flanking the posterior end of the cluster (Fig. 3).



[Figure 3]

Fig.3 : Elastic spring expansion under a weak force F1 (schematic**). A)** normal case: an elastic spring fastened completely (3 binding sticks) at its left end and pulled at its right end. Spot 1 is shifted to the right (corresponding to gene Hox1 entering in the TF domain). B) The spring is loosely fastened (1 binding stick) and it slides further to the right (corresponding to both Hox1 and Hox2 entering in TF. C) The elastic spring is completely loose (no binding stick) and all three spots move freely to the right (corresponding to all three Hox genes enter in TF) where they are expressed.

A)The traditional approach

As mentioned in the Introduction, for several decades the experimental tools and methods in use to explore the genetics of gene clusters were the chemical analysis of the biomolecules. This traditional methodology in Hox gene research is combined with genetic engineering techniques of DNA excision or duplication and the subsequent biomolecular analysis of the expression modifications of the neighboring Hox genes (see e.g. [19, 20]).

According to experiments performed with the above traditional techniques, it is established that the regulatory elements of the mouse Hox clusters are located upstream of the cluster even beyond gene Evx2. (Fig.1b). A detailed search has explored this upstream area. Partial excisions of this area cause variable hox gene expressions deviating from the normal (*wild type)* gene expressions [21].

The expanding spring approximation can explain the above excision results (ignoring for simplicity the local chromosomal interactions [18]). When a partial deletion of the regulatory region is performed, the fastening of the Hox cluster is incomplete so that the cluster is relatively loose and, under the action of a (weak) force **F1**, it can slide toward the telomeric side and gene Hox1 is expressed earlier that normally as shown schematically in Fig.3 [21]. For a more extended deletion of the regulatory region and under the action of the same force **F1,** the cluster slides further (Fig. 3). In the extreme case of a complete deletion of the DNA sequence even beyond Evx, the whole regulatory region will be excised and the Hox cluster will be loose and it can move like a free body. **F1** will shift the Hox cluster inside the activation region and no gradual gene expression can occur (Fig. 3).

The above expectation is a BM prediction in retrospect: amazingly, as early as 1999 T. Kondo and D. Duboule observed this phenomenon happening 'as if temporal collinearity disappeared' [21, p.414]. Therefore, BM directly relates the complete deletion of the regulatory elements of the HoxD cluster with the disappearance of TC as a result of the free movement of the whole cluster inside the gene activation domain.

B) A novel approach involving conserved non-coding elements (CNE)

The last twenty years, a novel powerful weapon has been added to the tools exploring the properties and activation of Hox gene clusters. Besides the protein coding genes in the genome, there are more than 30.000 RNA elements in the human genome which do not code any proteins (ncRNA). More than 1000 of these non-coding RNAs are persistently conserved from generation to generation [22]. Lately, the conserved DNA non-coding elements (CNE) and their long non-coding RNAs (lncRNA) have been intensively studied by many groups with the help of sophisticated numerical analysis methods [23,24]. Their findings are impressive. The numerous CNEs are preserved for more than 400 million years of evolution. The size of these CNEs varies, it can reach more than 300 bp. The CNEs play different roles in normal development and disease

coordinating spatial-temporal gene expression in both embryos and grown up animals. A CNE that has been thoroughly studied is *Hotair* which is a lncRNA located between HOXC11 and HOXC12 in the vertebrate chromosome 12 [24]. Its location in the posterior domain of the HOXC cluster is compatible with the hypothesis that CNEs can be involved in the creation of the pulling forces of BM. In what follows, CNEs play an important role particularly in the expanding spring approximation. It is further assumed that several CNEs are located in the fastening posterior domain of the vertebrate Hox clusters incorporating gene Evx2.

In the early studies of CNEs it was reported the striking differences in structure and function of the CNEs occur for various vertebrates, in particular in Humans and mice [25]. For instance the human *Hotair* represses *in trans* HOXD gene expressions while deletion of mouse Hotair *in vivo* does not affect the mouse Hoxd transcriptions. Later studies did not confirm that the deletion of mouse Hotair has a detectable effect on Hoxd gene expressions *in vivo* [26]. The situation is not clear since the mouse Hotair knock out causes derepression of HoxD genes [27]. Even if Hotair shows low sequence conservation in several vertebrates, it has been noticed that many ncRNAs are conserved in structure although not conserved in sequence [28].

In view of the above findings, an International Bioinformatics Group in collaboration with the Imperial College team headed by B. Lenhard have recently performed a comprehensive analysis of the regulatory roles of Hotair *in cis* and *in trans* on Hox clusters [24]. Nepal et al propose that at the second round of whole-genome duplication, HOTAIR expression is correlated positively with HOXC11 *in cis* and negatively correlated with HOXD11 *in trans* [24]. They compared human and zebrafish CNEs and identified a 32-nucleotide long CNE conserved across the vertebrates. They characterized this long CNE as the ancestral sequence of the ancestral HoxC/D cluster . Their conclusion is that a lncRNA locus functions at the DNA and RNA levels regulating genes both *in cis* and *in trans* [24] .This is a challenging hypothesis to be further tested. More specifically, this analysis indicates that HOTAIR expression regulates positively human HOXC11 in *cis* and negatively HOXD11 in *trans* [24].The number of the conserved elements varies, depending on the cluster copies during the whole genome duplications. Therefore, a certain pulling force will cause a variable shift (sliding) of the cluster toward its telomeric end and the starting time of Hox gene expression will be accordingly modified.

In the expanding spring approximation a Hox cluster is fastened at the neighboring posterior domain of the cluster [16,17]. This fastening is achieved with the action of a set of CNEs- an alternative approach from the traditional approach. The two versions lead to equivalent results.

As stated in the Introduction, in the case of *Xenopus laevis* M. Kondo et al. analyzed the Hox gene expressions and concluded that the initiation of these expressions deviates from the normal paralogue order of Hox gene expressions, therefore they violate TC [5]. Following the WGD the number of CNEs fastening the Hox clusters may vary so that the total effect on Hox cluster expansion differs for the various clusters. According to approach B, the initiation and duration of gene activation is reshuffled. This could explain the violation of TC observed in the L and S homologs of the *Xenopus leavis* as observed by M. Kondo et al. [5].

4. Conclusion

Some developmental models setup an abstract theoretical framework for the guiding principles like the positional information theory of L. Wolpert [29, 30]. In contrast, many other models are proposed which deal with more concrete phenomena equally important. A phenomenon of this category is the activation of Hox gene clusters.

In Section 3 two distinct approaches are presented for these gene expressions- traditional and novel approaches. The formulations, although looking different, they produce admissible results (not contradicting to experimental evidence). Eventually, this could make the two approaches equivalent. It is important to explore this possibility by trying to formulate (if possible) an experimental setup where the two approaches predict distinct results. The experimental confirmation of one of them would discard the other(s).

It has been suggested that TC is the 'principal constraining force' keeping Hox clusters in a compact organization [31]. In the case of a complete deletion of the regulatory region, BM predicts that the Hox cluster becomes completely loose and freely moving body. Its early slide toward the activation domain causes the disappearance of TC (Section 3).

The above BM prediction leads to a daring hypothesis for the absence of temporal collinearity in Drosophila. The mechanism of TC disappearance in Drosophila could be the same mechanism

predicted by BM combined with the evolutionary destruction pathway from an ancestral Hox cluster of organized type (O) to a Hox cluster lacking TC [4, 31].

References

1. Lewis E.B. A gene complex controlling segmentation in *Drosophila*. Nature. 1978;276:565–570. doi: 10.1038/276565a0. [PubMed] [CrossRef] [Google Scholar]

2. Sasaki H., Hogan B.L. Differential expression of multiple fork head related genes during gastrulation and axial pattern formation in the mouse embryo. Development. 1993;118:47–59. [PubMed] [Google Scholar]

3. Izpisua-Belmonte J.C., Duboule D. Murine genes related to the Drosophila AbdB homeotic genes are sequentially expressed. EMBO J. 1991:10:2279-2289.

4. Duboule D. The rise and fall of Hox gene clusters. Development. 2007;134:2549–2560. doi: 10.1242/dev.001065. [PubMed] [CrossRef] [Google Scholar]

5. Kondo M., Yamamoto T., Takahashi S., Taira M., Comprehensive analyses of Hox gene expression in *Xenopus laevis* embryos and adult tissues. Develop. Growth Differ. 2017:59:526-539.

6. Zhu K., Spaink H.P., Durston A.J. Collinear Hox-Hox interactions are involved in patterning the vertebrate anteroposterior (A-P) axis. PLoS ONE. 2017;11:e0175287. doi: 10.1371/journal.pone.0175287. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

7. Durston A.J. Vertebrate Hox temporal collinearity: does it exist and what is its function? Cell Cycle 2019:18:523-530.

8. Papageorgiou S. A physical force may expose Hox genes to express in a morphogenetic density gradient. Bull. Math. Biol. 2001;63:185–200. doi: 10.1006/bulm.2000.0211. [PubMed] [CrossRef] [Google Scholar]

9. Papageorgiou S. Pulling forces acting on Hox gene clusters cause expression collinearity. Int. J. Dev. Biol. 2006;50:301–308. doi: 10.1387/ijdb.052034sp. [PubMed] [CrossRef] [Google Scholar]

10. Papageorgiou S. Physical forces may cause Hox gene collinearity in the primary and secondary axes of the developing vertebrates. Dev. Growth Differ. 2011;53:1–8. doi: 10.1111/j.1440-169X.2010.01218.x. [PubMed] [CrossRef] [Google Scholar]

11. Papageorgiou S. A biophysical mechanism may control the collinearity of Hoxd genes during the early phase of limb development. Hum. Genomics. 2009;3:275–280. doi: 10.1186/1479-7364-3-3-275. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

12. Noordermeer D., Leleu M., Schorderet P., Joye E., Schabaud F., Duboule D. Temporal dynamics and developmental memory of 3D chromatin architecture at Hox gene loci. eLIFE 2014, 3.

13. Fabre P.J., Benke A., Joye E., Huynh T.H.N., Manley S., Duboule D. Nanoscale spatial organization of the HoxD gene cluster in distinct transcriptional states. Proc. Nat. Acad. Sci. 2015: 112:13964-13969.

14. Fabre P.J., Benke A., Manley S., Duboule D. Visualizing HoxD gene cluster at the nanoscale level. Cold Spring Harb. Symp. Quant. Biol. 2015;80:9–16. doi: 10.1101/sqb.2015.80.027177. [PubMed] [CrossRef] [Google Scholar]

15. Papageorgiou S. Comparison of models for the collinearity of Hox genes in the developmental axes of vertebrates. Curr.Genomics 2012: 13:245-251.

16. Papageorgiou S. Physical forces may cause the HoxD gene cluster elongation. Biology 2017:6:32.

17. Papageorgiou S. Abnormal elongations of HOX Gene clusters may cause cancer. Frontiers in Cell and Developmental Biology 2018: 6:25.

18. Mallo M., Alonso C.R. The regulation of Hox gene expression during animal development. Development 2013:140:3951-3963.

19. Tarchini B., Duboule D. Control of Hoxd genes' collinearity during early limb development. Dev. Cell. 2006;10:93–103. doi: 10.1016/j.devcel.2005.11.014. [PubMed] [CrossRef] [Google Scholar]

20. Parker H.J., Bronner M.E., Krumlauf R. The vertebrate Hox gene network for hindbrain segmentation. Bioessays. 2016;38:526–538. doi: 10.1002/bies.201600010. [PubMed] [CrossRef] [Google Scholar]

21. Kondo T., Duboule D. Breaking collinearity in the mouse HoxD complex. Cell. 1999;97:407–417. doi: 10.1016/S0092-8674(00)80749-7. [PubMed] [CrossRef] [Google Scholar]

22. Washieti S., Hofacker I.L., Lukasser M., Huttenhofer A., Stadler P.F. Mapping of conserved RNA secondary structures predicts thousands of functional noncoding RNAs in the human genome. Nature Biotechnology. 2005: 23:1383-1390.

23. Polychronopoulos D., King J.W.D., Nash A.J., Tan G., Lenhard B. Conserved non-coding elements: developmental gene regulation meets genome organization. Nucleic Acids Research. 2017: 45:22:12611-12624.

24. Nepal C., Taranta A., Hadzhiv Y., Pundhir S., Mydel P., Lenhard B., Muller F., Andersen J.B. Ancestrally duplicated conserved noncoding element suggests regulatory roles of HOTAIR in *cis* and *trans*. iScience. 2020: 23(4):101008.

25. Schorderet P., Duboule D. Structural and functional differences in the long non-coding RNA *Hotair*in mouse and human. PLoS Genetics, 2011: 7:5 e1002071.

26. Amandio A.R., Necsulea A., Joye E., Mascrez B., Duboule D. *Hotair* is dispensible for mouse development. PLOS Genetics. 2016. 12:12.

27. He S., Liu S., Zhu H. The sequence, structure and evolutionary features of HOTAIR in mammals. BMC Evol. Biol. 2011. 11:102.

28. Li L., Liu B., Wapinski O.L. *et al.* Targeted disruption of *Hotair* leads to homeotic transformation and gene –repression. *Cell Rep.* 2013. 5(1):3-12.

29. Durston A.J. A tribute to Lewis Wolpert and his ideas on the 50th Anniversary of his paper 'Positional information and the spatial pattern of differention'. Int. J. Mol.Sci. 2020. 21(7):2552.

30. Gordon N,K., Chen Z., Gordon R., Zou Y. French flag gradients and Turing reaction-diffusion theory . Biosystems, 2020; 196:104169

31. Garstang M., Ferrier D.E.K. Time is of the essence for ParaHox homeobox gene clustering. BMC Biol. 2013;11:72–75. doi: 10.1186/1741-7007-11-72. [PMC free article] [PubMed] [CrossRef] [Google Scholar]