


Review

# Co-Regulated Genes and Gene Clusters

Sergey V. Razin <sup>1,2,\*</sup> , Elena S. Ioudinkova <sup>1</sup>, Omar L. Kantidze <sup>1</sup> and Olga V. Iarovaia <sup>1</sup>

<sup>1</sup> Institute of Gene Biology Russian Academy of Sciences, 119334 Moscow, Russia; ioudinkova@inbox.ru (E.S.I.); o\_kantidze@mail.ru (O.L.K.); iarovaia@inbox.ru (O.V.I.)

<sup>2</sup> Faculty of Biology, M.V. Lomonosov Moscow State University, 119234 Moscow, Russia

\* Correspondence: sergey.v.razin@usa.net; Tel.: +7-499-135-97-87

**Abstract:** There are many co-regulated genes in eukaryotic cells. The coordinated activation or repression of such genes occurs at specific stages of differentiation, or under the influence of external stimuli. As a rule, co-regulated genes are dispersed in the genome. However, there are also gene clusters, which contain paralogous genes that encode proteins with similar functions. In this aspect, they differ significantly from bacterial operons containing functionally linked genes that are not paralogs. In this review, we discuss the reasons for the existence of gene clusters in vertebrate cells and propose that clustering is necessary to ensure the possibility of selective activation of one of several similar genes.

**Keywords:** co-regulated genes; gene clusters; coordinated expression



**Citation:** Razin, S.V.; Ioudinkova, E.S.; Kantidze, O.L.; Iarovaia, O.V. Co-Regulated Genes and Gene Clusters. *Genes* **2021**, *12*, 907. <https://doi.org/10.3390/genes12060907>

Academic Editor: Tatyana D. Kolesnikova

Received: 28 May 2021  
Accepted: 10 June 2021  
Published: 11 June 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In living systems, there is often a need for a coordinated change in the expression level of a group of genes. A classic example is the activation of genes encoding the enzymes required for utilizing an alternative energy source in unicellular organisms. In prokaryotic cells, genes are grouped into operons, which are transcribed as polycistronic mRNAs [1–3]. In the genomes of the vast majority of eukaryotic organisms, classical operons (i.e., clusters of functionally linked genes that are transcribed as part of polycistronic mRNA) are absent. The only exception is the *Caenorhabditis elegans* genome, where a significant part of the genes is grouped into operons of various types. The most common operons are the SL1 type, the polycistronic mRNA of which is divided into a set of individual mRNAs through trans-splicing [4,5]. Single operons of this type have also been found in some other organisms [6]. At the same time, for unicellular eukaryotic organisms (yeast [7,8]), fungi [9,10]), and plants [11–14]), the organization of functionally linked genes into clusters is quite typical. Within the clusters, each gene has its own promoter, but there are also some regulatory elements that control all genes in the cluster. Thus, in the gene clusters encoding the synthesis of secondary metabolites in fungi, there is often a gene encoding a transcription factor that activates the promoters of the remaining genes of the cluster [15]. Being a trans-acting factor, for some reason, this activator only works within the gene cluster. The transfer of a target gene from a cluster to an ectopic position elsewhere in the genome, results in the loss of the ability to be activated by a cluster-specific regulatory factor. Along with cluster-specific (metabolic pathway-specific) transcription factors in fungi, there are also global regulators that coordinate the work of different metabolic pathways [15]. In other words, there is a multilevel regulatory system that is also common in mammalian and other vertebrate cells.

In fungi and plants, gene clusters often contain functionally related genes that are not paralogs. The reasons for the emergence and stabilization of such gene clusters in evolution are the subject of scientific discussions [16,17]. In vertebrate animals, gene clusters typically arise as a result of gene duplications, with subsequent specialization of the functions of proteins encoded by individual genes—the globin, histocompatibility complex, Hox, and olfactory receptor genes are examples of such paralogous gene clusters.

The coordinated activation of a large number of genes in eukaryotic cells is typically associated with the differentiation process, although there are exceptions, such as stress-response gene activation. Genes activated during differentiation can be either organized into clusters (e.g., clusters of globin genes) or scattered throughout the genome. In this review, we will analyze the mechanisms that control the expression of tissue-specific genes and examine the advantages of organizing such genes into clusters. The discussion will be largely based on the studies of vertebrate erythroid-specific genes. However, as necessary, we will refer to other tissue-specific genes and gene clusters.

## 2. Erythroid-Specific Genes

In the course of erythroid differentiation, transcription of a significant number of genes is activated, whereas many other genes are repressed. The final aim of erythroid differentiation is the generation of hemoglobin-producing cells. In vertebrates, hemoglobin comprises 4 polypeptide chains (two alpha- and two beta-type globin chains) and heme. In warm-blooded animals, genes encoding for the alpha and beta subunits of hemoglobin are segregated into two clusters, each containing several genes that are expressed at different developmental stages [18]. In fish and amphibians, the alpha- and beta-globin genes are located in a common cluster (or common clusters) [19–21]. However, these clusters also contain groups of genes that are specifically expressed in particular developmental stages. For many years, the mammalian and avian clusters of alpha- and beta-globin genes have served as the most popular models for studying the regulation of gene expression [22,23]. Each of these clusters is controlled by a super-enhancer, which in the beta- and alpha-globin gene clusters is referred to as the locus control region (LCR) [24] and major regulatory element (MRE), respectively [25]. In addition to the super-enhancer, a number of additional enhancers are involved in the control of globin gene expression. These enhancers can be located both within gene clusters and at a considerable distance from these clusters [26,27]. In globin-producing cells, all these enhancers are assembled into a single block, to which various globin genes are recruited in a stage-specific fashion [28–31].

The activity of enhancers is controlled by a set of erythroid-specific transcription factors (Gata1, Tal1, Klf1, and NFE2) [32]. The current model suggests that, being bound to an enhancer, these transcription factors recruit transcription machinery and trigger the assembly of a liquid activator compartment, through a physical process of liquid phase separation [33,34]. Transcriptional activation of individual globin genes correlates with the recruitment of their promoters to the activator compartment [30,31]. In accordance with this scenario, a number of studies have demonstrated that the same transcription factors are present both on the enhancer and on the promoters of genes controlled by this enhancer [35–37]. Due to the higher density of binding sites for transcription factors on enhancers, it is the enhancers that serve as the nucleation centers for the assembly of activator compartments [38–40]. The attraction of promoters to the activator compartment is controlled by architectural proteins. In erythroid cells, such architectural proteins are LMO2 and LDB1 [41].

There are significant differences in the transcription control mechanisms of the alpha- and beta-globin gene clusters [42]. A cluster of beta-globin genes is integrated into a more extended cluster of tissue-specific olfactory receptor genes. These genes are inactive in most cells. Accordingly, in the vast majority of non-erythroid cells, the extended genome region harboring both olfactory receptor genes and beta-globin genes is organized into inactive (DNase resistant) chromatin. In the course of erythroid differentiation, the chromatin status of the beta-globin gene cluster changes, as evidenced by an increase in the level of histone acetylation and the transition of the entire cluster to an active DNase-sensitive chromatin configuration [43,44]. This process is controlled by LCR and other enhancers located in the upstream region of the beta-globin gene cluster. The loss of this complex of enhancers as a result of the natural deletion of an extended genomic segment (Hispanic deletion) results in an inability to activate the beta-globin gene cluster in erythroid cells [45]. A cluster of alpha-globin genes is located in a permanently active region of the genome. It is surrounded by

housekeeping genes and is found in an active (DNase-sensitive) chromatin configuration, in both erythroid and non-erythroid cells [42]. The main regulatory element of the alpha-globin gene cluster is located in the intron of an upstream housekeeping gene [46]. At least in the mouse, alpha-globin genes are transcribed in mixed transcription factories, along with adjacent housekeeping genes [47]. The current model postulates that the expression of both alpha- and beta-globin genes in erythroid cells is controlled by mechanisms operating at two levels [26,48,49]. The first level is the activation of the super-enhancer and the assembly of the enhancer block. In a cluster of beta-globin genes, this activation correlates with a change in the chromatin configuration of the entire gene cluster (a transition from DNase-resistant to DNase-sensitive configuration, which in turn, is ensured by an increase in the level of histone acetylation [44,45]). Super-enhancer activation is mediated by several erythroid-specific transcription factors, the most important of which are Gata1, Tal1, Klf1, and NFE2 [50,51]. The same factors activate the expression of other erythroid-specific genes that are not organized into clusters. The coordination of the expression of all erythroid-specific genes is provided primarily by the ability of their promoters to be activated by the above-mentioned erythroid-specific transcription factors. The expression of the transcription factors themselves is also coordinated in a certain way. According to a number of data, an important role in this process is played by the establishment of spatial contacts between genes that code for transcription factors. Architectural proteins, primarily CTCF [52], are involved in establishing such contacts. The binding of CTCF to recognition sites on DNA, in turn, can be modulated by specific non-coding RNAs [53]. Spatial contacts are also established between downstream targets of the key erythroid-specific transcription factors. Along with other factors, the recruitment of erythroid-specific genes to specialized transcription factories [54] contributes to the generation of an erythroid-specific gene interactome.

### 3. Functional Role of Gene Clustering in Vertebrate Cells

As aforementioned, the organization of functionally related genes into operons in prokaryotes and operon-like clusters in fungi and plants, is typically associated with the necessity to ensure coordination in the production of components of one protein complex or enzymes of one metabolic pathway. It is quite clear that such a function cannot be attributed to the above-discussed clusters of alpha- and beta-globin genes. Indeed, at each stage of development, preferably only one of the genes of each cluster is expressed. Even in cases when two globin genes of the same type are expressed (for example, in mice,  $\beta$ major and  $\beta$ minor adult globin genes), the production of one of them ( $\beta$ major) exceeds several times the production of the other. An important task is coordinating the expression level of alpha- and beta-globin chains. However, in warm-blooded animals, the corresponding genes are segregated in different clusters located on different chromosomes, and the coordination of their expression is provided by regulatory mechanisms of the first level. How this coordination is achieved is not entirely clear. Some observations suggest that in erythroblasts, the genes encoding the alpha and beta chains of hemoglobin are spatially close to each other [55]. However, this relative spatial proximity is far from a colocalization that may be expected for the genes attracted to a common transcription factory. At best, alpha- and beta-globin genes are often located close to the same speckle (SC35 domain), but the significance of this observation is not clear.

As globin gene clustering is not important for the hemoglobin subunit expression level coordination, what could the role of such clustering be? To answer this question, it is worth remembering that at each developmental stage, only one alpha-type globin gene and one beta-type globin gene are preferentially expressed. In mammals, there are beta-globin genes of the embryonic, fetal, and adult types and alpha-globin genes of the embryonic and adult types. Ensuring the expression of one of several similar genes is a rather difficult task that cannot be solved by regulatory systems of the first level. Apparently, this difficulty is due to the fact that the regulatory modules in the promoters of all globin genes are quite similar, and all these genes are, in principle, activated by the same transcription factors. We

believe that precisely to make the choice of the gene that should work at a particular stage of development, second-level regulatory systems are required that work at the gene cluster level. The regulation is carried out in this case in cis, by creating alternative enhancer-promoter loops [30,31,56,57]. For this regulatory mechanism to work, it is important that the regulated genes are close to each other; that is, they are organized into a gene cluster. Within the cluster, the selection of genes to be expressed may be mediated by different mechanisms, including the ones that ensure selective repression of some genes of the cluster. The common theme is recruitment of co-repressors mediating histone deacetylation. In some cases, this is accompanied by DNA methylation [58–62].

How the spatial reconfiguration of globin gene clusters is directed is not entirely clear. There are many indications that an important role in this process is played by a change in the binding profile of CTCF, to recognition sites on DNA [52,63,64]. The interaction between CTCF binding sites is modulated by histone acetylation in the surrounding regions, which in turn is directed by erythroid-specific transcription factors, in particular GATA1 [65,66]. Non-coding RNAs can also influence the preference for CTCF binding to different recognition sites on DNA [53]. It is worth emphasizing once again that the correct switching of globin gene expression is possible only in the context of existing gene clusters. A simple change in the order of the genes, relative to the super-enhancer (for example, when the entire gene cluster is inverted), as well as the inclusion of an additional gene in the cluster, leads to serious disruptions in the switching process of both alpha- and beta-globin genes [67–69]. In experiments with ectopic expression, the adult beta-globin gene controlled by the LCR was expressed at an early stage of development, when this gene should not be expressed. The same gene in the construct, which additionally contained fetal-globin genes, showed correct stage-specific expression [70,71]. Apparently, the competition of different genes in the cluster for the LCR is essential for ensuring the stage-specific expression of individual genes.

The main conclusion that follows from the above discussion is that clustering of globin genes is necessary for the operation of mechanisms that ensure the selection of those genes from a number of paralogous genes that will be expressed in a particular situation. Similar trends can be traced in the case of other gene clusters of vertebrates. Hox gene clusters are a good example. In these clusters, the activation time of individual genes during development is determined by the order of the genes in the cluster [72,73]. As is the case for globin genes, Hox genes are subject to multilevel regulation through global regulatory elements located outside the gene clusters and additional enhancers [74–76]. Activation of various groups of Hox genes is associated with spatial reconfiguration of gene clusters; that is, it is provided by mechanisms that work in cis [77–80].

Odorant receptor genes are another example of clustered vertebrate genes. In humans and mice genomes, there are several clusters of odorant receptor genes located on different chromosomes. In the course of differentiation, transcription of only one of the odorant receptor genes is activated in an individual olfactory neuron. In the mouse genome, cluster-specific regulatory elements were found, which provided random activation of one of the genes in the cluster [81–83]. How coordination between different clusters is ensured, resulting in the expression of only the gene activated in one of the clusters, remains not entirely clear. According to a number of data, inactive genes from different chromosomes are combined in the repressive compartment [84,85]. From the standpoint of our discussion, however, it is important that in the case of olfactory receptor genes, clustering is required to select one of the genes with similar regulatory modules, which is determined by the cluster-specific regulatory elements [81–83,86]. A number of other examples of gene clusters can be cited, where out of several similar genes, only one is selected to be expressed. Thus, this process occurs in the human growth hormone (hGH) multigene cluster [87]. The same situation can be observed in clusters of protocadherin (*Pcdh*) genes. The mammalian genome contains three clusters of *Pcdh* genes (*Pcdh $\alpha$* , *Pcdh $\beta$* , and *Pcdh $\gamma$* ). *Pcdh $\alpha$*  and *Pcdh $\gamma$*  clusters are constructed in a similar manner—tandem arrays of genes encoding the first variable exon and then a group of three small constant exons. This genomic architecture is

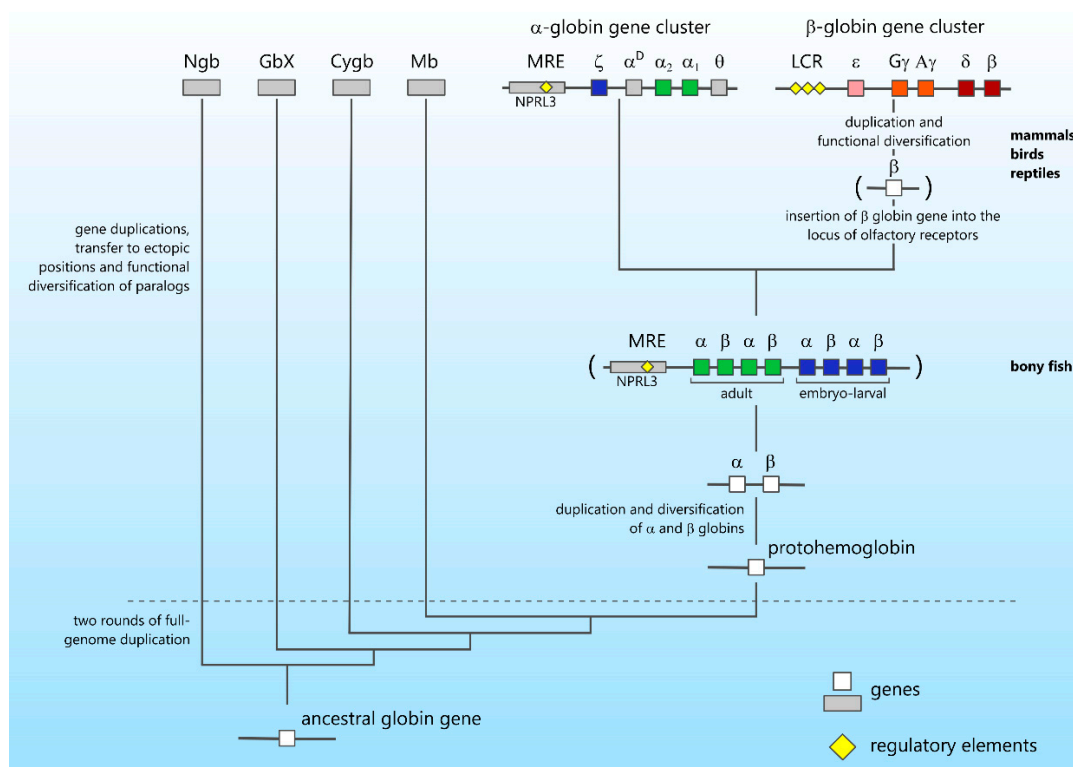
similar to that of immunoglobulin and the T cell receptor gene clusters. Within the *Pcdha* and *Pcdhγ* clusters, each of the genes encoding the variable exon has its own promoter. The transcript of a variable exon initiated on this promoter is spliced to a common set of three downstream small constant exons, within the respective cluster. Hence, through stochastic activation of the promoter and cis-alternative splicing, clustered *Pcdhs* can generate dozens of different isoforms. The *Pcdhβ* gene cluster contains only variable genes, each of which encodes a full-length protein. The *Pcdha* cluster is regulated by a 3' super-enhancer consisting of two cis-regulatory elements. The expression of *Pcdhβ* and *Pcdhγ* clusters is controlled by a super-enhancer located in the 3'-region of the *Pcdhγ* cluster [88]. The assembly of enhancer-promoter loops is mediated by CTCF, the binding sites of which are present in all promoters and enhancers, and in the *Pcdha* cluster, the promoter of each alternative gene is flanked by two CTCF binding sites (an additional one is located in the coding region). Enhancer-promoter loops are formed between forward and reverse CTCF-binding sites via CTCF/cohesin-mediated chromatin loop extrusion. Chromatin loop formation is regulated epigenetically through the methylation of CpG dinucleotides at the CTCF binding sites on the *Pcdhs*' promoters [89]. Accordingly, demethylation of the CTCF-binding site at a particular promoter allows the interaction of this promoter with the enhancer, resulting in activation of expression of this particular *Pcdha* gene [90]. Recent studies have shown that the aforementioned additional CTCF site of each alternative exon is associated with an antisense promoter that controls the transcription of long non-coding RNA (lncRNA). Stochastic transcription of this RNA through the sense promoter, triggers DNA demethylation of the corresponding CTCF-binding site, allowing CTCF binding and subsequent activation of the sense promoter [90]. The selection of the *Pcdh* promoter via the establishment of specific patterns of CTCF binding sites' methylation across the gene cluster, occurs early in the course of embryonic stem cell differentiation, and is then stably inherited by differentiated neurons [91]. Bioinformatic analysis also revealed several other gene clusters, organized similarly to *Pcdh* gene clusters [92]. Clearly, in all these gene clusters, the clustering of similar genes is necessary to select one from many genes, which will be expressed in a particular cell. In globin and *Pcdh* gene clusters, this selection is ensured by establishing alternative 3D cluster configurations. However, other mechanisms are also possible. Thus, immunoglobulin gene clusters are rearranged at the DNA level [93].

Regarding tissue-specific genes not being organized into clusters, their expression can be controlled both by the regulatory systems of the first level (direct activation of promoters by tissue-specific transcription factors) and in two stages—activation of a tissue-specific enhancer and subsequent activation of secondary targets. The genome distribution of genes not organized in dense clusters is not entirely random [94–96]. Several observations suggest that they can form sparse clusters corresponding to the structural and functional blocks of the genome (such as TADs or insulated neighborhoods), which limit the areas of tissue-specific enhancer activity [97–99]. However, genome-wide analysis did not reveal a good match between TADs and the co-expression domains [100].

#### 4. Gene Cluster Evolution

The evolution of gene clusters can be traced by the example of clustered globin genes (Figure 1). Several interesting reviews have been devoted to the role of duplications in globin gene evolution [101,102]. The ancestor of globin genes is believed to simultaneously appear in the eukaryotic genome with mitochondria and plastids, as a result of gene transfer or genome fusion [103]. Most likely, the functions of the ancient membrane-bound globin were related to signaling or protection [104]. The family of globin genes arose as a result of neo- and sub-functionalization of additional copies of genes that arose after whole-genome and segmental duplications, which repeatedly occurred during the evolutionary process [105–107]. In the genomes of modern vertebrates, there are eight groups of globin genes—neuroglobin (*Ngb*), cytoglobin (*Cygb*), androglobin (*Adgb*), hemoglobin (*Hb*), and myoglobin (*Mb*), which are present in almost all vertebrate genomes, as well as the less common globins X, Y, and E (*GbX*, *GbY*, and *GbE*) [106,107]. Protoneuroglobin is believed

to be the common ancestor of the entire family of globin proteins in both protostomes and deuterostomes [108]. Gene duplications have played an exceptional role in vertebrate evolution. The two rounds of whole-genome duplications (WGDs) that occurred at the dawn of vertebrate's radiation provided fuel for the evolutionary process, and made possible the emergence of fundamentally new traits and regulatory systems [109–116]. As a result of two WGD cycles, specialized globins that enabled the storage (Mb and Cygb) and transport (Hb) of oxygen appeared in the common ancestor of vertebrates. The specialization and functional diversification of globin genes was related to both genome-wide and tandem duplications, which repeatedly occurred with globin genes [117,118]. Tandem duplication, which resulted in the formation of alpha- and beta-globin genes, occurred 450–500 million years ago in the Ordovician, before the separation of cartilaginous fish and the common ancestor of modern tetrapods [117,119]. It is assumed that the genomic context and structure of the postduplication ancestral locus was similar to the structure of the major locus of alpha-/beta-globin genes of modern teleost fishes [20]. Within the locus, alpha- and beta-globin genes alternate and are transcribed in opposite directions. The organization of alpha- and beta-globin genes into a common cluster controlled by an upstream regulatory element that is typical for fish is also preserved in amphibians [120–123]. It should be noted that the emergence of clustered globin genes in the course of evolution correlates with the diversification of these genes in terms of expression at particular developmental stages. For example, in the *Danio rerio* genome, the major locus of globin genes is structurally and functionally segregated into two stage-specific subdomains, which contain genes that are expressed during the adult and embryonic-larval stages of development [19,124].



**Figure 1.** Main globin gene evolution stages. Clusters of human alpha- and beta-globin genes are shown at the top of the evolution tree. In the alpha-globin gene cluster, the colors of embryonal and adult globin genes show that these genes are homologs of embryonal and adult globin genes present in the ancestral cluster. The different color scheme for designation of embryonal ( $\epsilon$ ), fetal ( $G\gamma$ ,  $A\gamma$ ), and adult ( $\delta$ ,  $\beta$ ) globin genes in the beta-globin gene cluster is used to show that these genes are not homologs of embryonal and adult globin genes present in the ancestral cluster, but originated via duplication of a single beta-type globin gene transposed to a new genomic position.

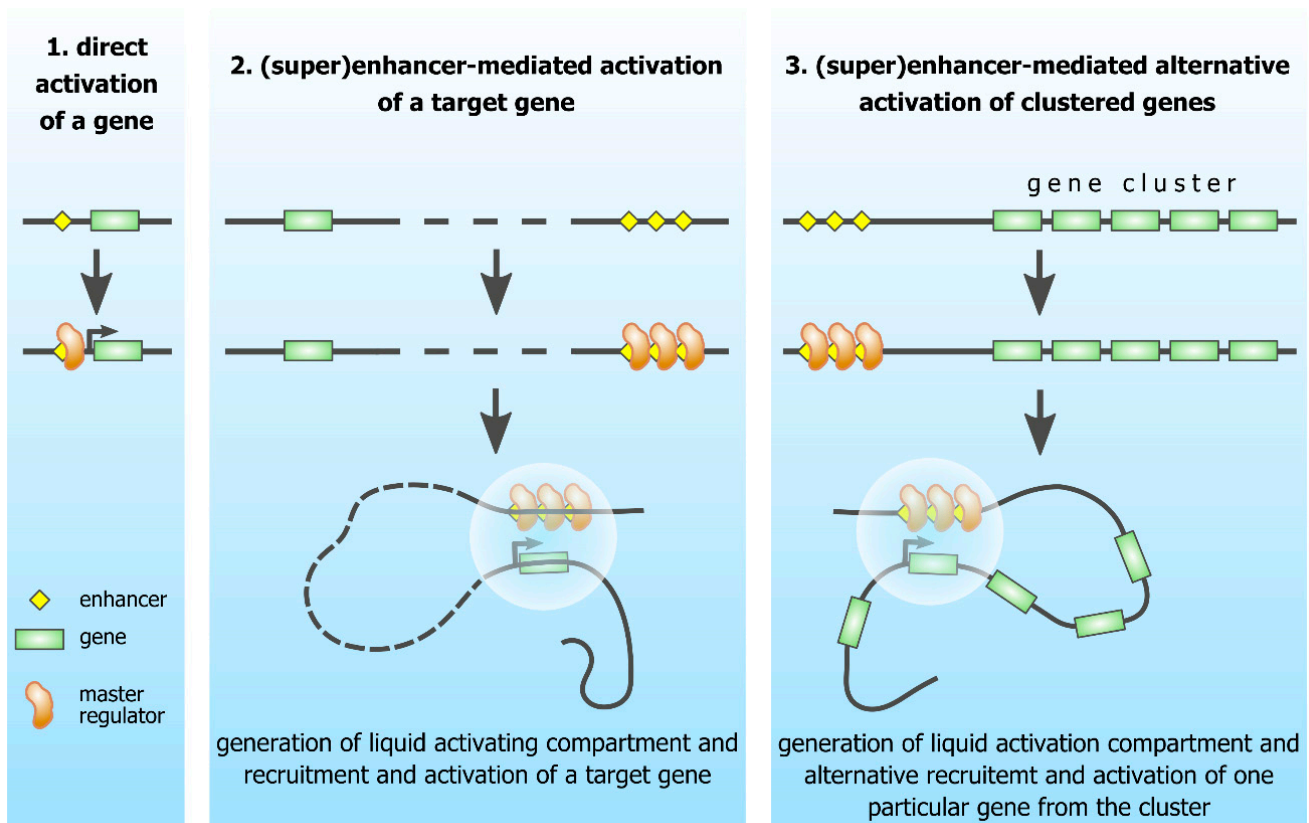
The next important stage in the evolution of globin genes was the segregation of alpha- and beta-globin genes in amniotes. The key event in this case was the relocation of a part of the ancestral cluster or, more likely, one gene pair to a new genomic position (into the odorant receptor gene cluster) [20,120,125,126]. The subsequent loss of the alpha-globin gene(s) and duplications of the ancestral beta-globin gene gave rise to modern clusters of beta-globin genes. It should be mentioned that the specialization of embryonic and adult genes within the alpha-globin gene cluster occurred within the ancestral domain [127], whereas the emergence of stage-specific forms in the domain of beta-globin genes occurred as a result of a set of duplications after the transfer of the ancestral beta-globin gene to another chromosome [128] (see schematic in Figure 1). Duplications of beta-globin genes occurred independently in mammals and birds (i.e., stage-specific forms of hemoglobin were invented several times during evolution) [128,129]. It should be mentioned that gene duplications occur relatively often. However, in most cases, one of the duplicated copies is lost in the course of subsequent evolution (turns into a pseudogene). For example, the platypus alpha-globin gene cluster contains the non-functional beta-type omega globin gene [125,130]. The diversification of globin functions into embryonic, fetal (primate-specific), and adult ones contributed to the preservation of duplicated globin genes, which resulted in the emergence of the gene cluster [18,128].

As already mentioned, gene cluster integrity is important for the normal expression of the genes contained in the cluster. Moving one or more genes of the cluster to an ectopic position excludes these genes from the regulatory chain. However, such genes retain tissue-specific promoters, which, under certain circumstances, can be integrated into another regulatory network. As aforementioned, gene duplicates are often lost in the course of evolution. However, moving them to an ectopic genomic position may contribute to their preservation and neofunctionalization, in accordance with the new genomic context [131]. Apparently, this process contributed to the diversification of the functions of neuroglobin and other globins, as well as other non-clustered globin genes. In this regard, it is also interesting to trace what happens to foreign genes, which were moved to the locus of tissue-specific genes, due to genomic rearrangement. A cluster of tissue-specific genes is a single regulatory domain under the control of the first-level regulatory systems that direct tissue-specific expression. An alien gene trapped in this regulatory network most likely will not be able to perform its original function. However, there is a possibility that the promoter of this gene will eventually be adapted to work in the regulatory context of the gene cluster. An example of such a scenario is the chicken *TMEM8* gene, which was relocated downstream of the alpha-globin cluster, as a result of chromosomal inversion. Being relocated to the domain of chicken alpha-globin genes, this gene encoding a transmembrane protein was included in the regulatory network of this domain and acquired an erythroid-specific expression profile [132]. A similar scenario can be traced in the case of the folate receptor gene, which in the chicken genome, is located close to the cluster of beta-globin genes. Although this gene is regulated independently of the genes of the beta-globin cluster, it nevertheless acquired an erythroid-specific expression profile [133].

## 5. Concluding Remarks

In sum, we can state that in the genomes of the vast majority of eukaryotes, the coordinated expression of a number of genes (for example, genes encoding enzymes of one metabolic pathway, genes for which the expression begins in response to hormone action, or tissue-specific genes in the course of differentiation) is not conjugated with the organization of these genes into functional clusters. This coordination is provided by specific transcription factors that either directly interact with the promoters of the activated genes or activate a specific set of enhancers, to which, in turn, the genes dependent on these enhancers are recruited (Figure 2). The loss of operons in the genomes of most eukaryotic organisms is apparently associated with a significant increase in the number of co-regulated genes, and the need to activate different combinations of genes in response

to specific stimuli. In this situation, clusters of functionally linked genes are useless. At the same time, clustering turned out to be in demand for solving a different biological problem. In eukaryotic genomes, gene clusters typically contain paralogous genes resulting from tandem duplications. In the course of evolution, there has been a certain functional diversification of paralogous genes associated with their adaptation, to perform similar but still different functions. Accordingly, in specific situations, it becomes necessary to activate the expression of only one or several similar genes. It is this task that is solved at the level of gene clusters, via spatial (and sometimes genomic) reconfiguration of the gene cluster, which ensures the selective attraction of one of the genes to the enhancer (Figure 2).



**Figure 2.** Strategies of transcriptional control of dispersed and clustered tissue-specific genes.

**Author Contributions:** Conceptualization, S.V.R., O.V.I. and E.S.I.; writing—original draft preparation: S.V.R., O.V.I. and E.S.I.; writing—review and editing, S.V.R., O.V.I., E.S.I. and O.L.K.; figure visualization, O.L.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by Russian Science Foundation (RSF), grant number 21-64-00001 and by Russian Foundation for Basic Research, grant number 20-04-00003.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Liu, Z.; Feng, J.; Yu, B.; Ma, Q.; Liu, B. The functional determinants in the organization of bacterial genomes. *Brief. Bioinform.* **2021**, *22*. [[CrossRef](#)] [[PubMed](#)]
2. Reznikoff, W.S. The Operon Revisited. *Annu. Rev. Genet.* **1972**, *6*, 133–156. [[CrossRef](#)] [[PubMed](#)]
3. Jacob, F.; Monod, J. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **1961**, *3*, 318–356. [[CrossRef](#)]
4. Blumenthal, T. Operon and non-operon gene clusters in the *C. elegans* genome. *WormBook* **2014**, 1–20. [[CrossRef](#)] [[PubMed](#)]



5. Blumenthal, T.; Evans, N.; Link, C.D.; Guffanti, A.; Lawson, D.; Thierry-Mieg, J.; Thierry-Mieg, D.; Chiu, W.L.; Duke, K.; Kiraly, M.; et al. A global analysis of *Caenorhabditis elegans* operons. *Nature* **2002**, *417*, 851–854. [[CrossRef](#)] [[PubMed](#)]
6. Lasda, E.L.; Blumenthal, T. Trans-splicing. *Wiley Interdiscip. Rev. RNA* **2011**, *2*, 417–434. [[CrossRef](#)]
7. Wong, S.; Wolfe, K.H. Birth of a metabolic gene cluster in yeast by adaptive gene relocation. *Nat. Genet.* **2005**, *37*, 777–782. [[CrossRef](#)] [[PubMed](#)]
8. Hittinger, C.T.; Rokas, A.; Carroll, S.B. Parallel inactivation of multiple GAL pathway genes and ecological diversification in yeasts. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 14144–14149. [[CrossRef](#)] [[PubMed](#)]
9. Brakhage, A.A. Regulation of fungal secondary metabolism. *Nat. Rev. Genet.* **2012**, *11*, 21–32. [[CrossRef](#)]
10. Keller, N.P.; Turner, G.; Bennett, J.W. Fungal secondary metabolism from biochemistry to genomics. *Nat. Rev. Genet.* **2005**, *3*, 937–947. [[CrossRef](#)] [[PubMed](#)]
11. Frey, M.; Chomet, P.; Glawischnig, E.; Stettner, C.; Grün, S.; Winklmaier, A.; Eisenreich, W.; Bacher, A.; Meeley, R.B.; Briggs, S.P.; et al. Analysis of a Chemical Plant Defense Mechanism in Grasses. *Science* **1997**, *277*, 696–699. [[CrossRef](#)] [[PubMed](#)]
12. Kitaoka, N.; Zhang, J.; Oyagbenro, R.K.; Brown, B.; Wu, Y.; Yang, B.; Li, Z.; Peters, R.J. Interdependent evolution of biosynthetic gene clusters for momilactone production in rice. *Plant Cell* **2021**, *33*, 290–305. [[CrossRef](#)] [[PubMed](#)]
13. Mao, L.; Kawaide, H.; Higuchi, T.; Chen, M.; Miyamoto, K.; Hirata, Y.; Kimura, H.; Miyazaki, S.; Teruya, M.; Fujiwara, K.; et al. Genomic evidence for convergent evolution of gene clusters for momilactone biosynthesis in land plants. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 12472–12480. [[CrossRef](#)]
14. Boycheva, S.; Daviet, L.; Wolfender, J.-L.; Fitzpatrick, T.B. The rise of operon-like gene clusters in plants. *Trends Plant Sci.* **2014**, *19*, 447–459. [[CrossRef](#)] [[PubMed](#)]
15. Collemare, J.; Seidl, M.F. Chromatin-dependent regulation of secondary metabolite biosynthesis in fungi: Is the picture complete? *FEMS Microbiol. Rev.* **2019**, *43*, 591–607. [[CrossRef](#)] [[PubMed](#)]
16. Hurst, L.D.; Pál, C.; Lercher, M.J. The evolutionary dynamics of eukaryotic gene order. *Nat. Rev. Genet.* **2004**, *5*, 299–310. [[CrossRef](#)]
17. Nützmann, H.-W.; Scazzocchio, C.; Osbourn, A. Metabolic Gene Clusters in Eukaryotes. *Annu. Rev. Genet.* **2018**, *52*, 159–183. [[CrossRef](#)] [[PubMed](#)]
18. Philipsen, S.; Hardison, R.C. Evolution of hemoglobin loci and their regulatory elements. *Blood Cells Mol. Dis.* **2018**, *70*, 2–12. [[CrossRef](#)]
19. Ganis, J.J.; Hsia, N.; Trompouki, E.; De Jong, J.L.; DiBiase, A.; Lambert, J.S.; Jia, Z.; Sabo, P.J.; Weaver, M.; Sandstrom, R.; et al. Zebrafish globin switching occurs in two developmental stages and is controlled by the LCR. *Dev. Biol.* **2012**, *366*, 185–194. [[CrossRef](#)] [[PubMed](#)]
20. Hardison, R.C. Evolution of Hemoglobin and Its Genes. *Cold Spring Harb. Perspect. Med.* **2012**, *2*, a011627. [[CrossRef](#)] [[PubMed](#)]
21. Fuchs, C.; Burmester, T.; Hankeln, T. The amphibian globin gene repertoire as revealed by the *Xenopus* genome. *Cytogenet. Genome Res.* **2006**, *112*, 296–306. [[CrossRef](#)] [[PubMed](#)]
22. Razin, S.V.; Recillas-Targa, F. Chromatin domains and regulation of gene expression: Familiar and enigmatic clusters of chicken globin genes. *Crit. Rev. Eukaryot. Gene Expr.* **2001**, *11*, 16. [[CrossRef](#)]
23. Razin, S.V.; Farrell, C.M.; Recillas-Targa, F. Genomic Domains and Regulatory Elements Operating at the Domain Level. *Int. Rev. Cytol.* **2003**, *226*, 63–125. [[CrossRef](#)] [[PubMed](#)]
24. Grosveld, F.; van Assandelt, G.B.; Greaves, D.R.; Kollias, B. Position-independent, high-level expression of the human  $\beta$ -globin gene in transgenic mice. *Cell* **1987**, *51*, 975–985. [[CrossRef](#)]
25. Jarman, A.P.; Wood, W.G.; Sharpe, J.A.; Gourdon, G.; Ayyub, H.; Higgs, D.R. Characterization of the major regulatory element upstream of the human  $\alpha$ -globin gene cluster. *Mol. Cell. Biol.* **1991**, *11*, 4679–4689. [[CrossRef](#)] [[PubMed](#)]
26. Oudelaar, A.M.; A Beagrie, R.; Kassouf, M.T.; Higgs, D.R. The mouse  $\alpha$ -globin cluster: A paradigm for studying genome regulation and organization. *Curr. Opin. Genet. Dev.* **2021**, *67*, 18–24. [[CrossRef](#)] [[PubMed](#)]
27. Tolhuis, B.; Palstra, R.-J.; Splinter, E.; Grosveld, F.; de Laat, W. Looping and Interaction between Hypersensitive Sites in the Active  $\beta$ -globin Locus. *Mol. Cell* **2002**, *10*, 1453–1465. [[CrossRef](#)]
28. Vernimmen, D.; De Gobbi, M.; A Sloane-Stanley, J.; Wood, W.G.; Higgs, D.R. Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. *EMBO J.* **2007**, *26*, 2041–2051. [[CrossRef](#)]
29. Vernimmen, D.; Marques-Kranc, F.; Sharpe, J.A.; Sloane-Stanley, J.A.; Wood, W.G.; Wallace, H.A.C.; Smith, A.J.H.; Higgs, D.R. Chromosome looping at the human  $\alpha$ -globin locus is mediated via the major upstream regulatory element (HS –40). *Blood* **2009**, *114*, 4253–4260. [[CrossRef](#)]
30. De Laat, W.; Grosveld, F. Spatial organization of gene expression: The active chromatin hub. *Chromosome Res.* **2003**, *11*, 447–459. [[CrossRef](#)]
31. De Laat, W.; Klous, P.; Kooren, J.; Noordermeer, D.; Palstra, R.J.; Simonis, M.; Splinter, E.; Grosveld, F. Three-dimensional organization of gene expression in erythroid cells. *Curr. Top. Dev. Biol.* **2008**, *82*, 117–139. [[PubMed](#)]
32. Cantor, A.B.; Orkin, S.H. Transcriptional regulation of erythropoiesis: An affair involving multiple partners. *Oncogene* **2002**, *21*, 3368–3376. [[CrossRef](#)]
33. Gurumurthy, A.; Shen, Y.; Gunn, E.M.; Bungert, J. Phase Separation and Transcription Regulation: Are Super-Enhancers and Locus Control Regions Primary Sites of Transcription Complex Assembly? *BioEssays* **2018**, *41*, e1800164. [[CrossRef](#)] [[PubMed](#)]

34. Wang, X.; Cairns, M.J.; Yan, J. Super-enhancers in transcriptional regulation and genome organization. *Nucleic Acids Res.* **2019**, *47*, 11481–11496. [[CrossRef](#)]
35. Cho, Y.; Song, S.-H.; Lee, J.J.; Choi, N.; Kim, C.G.; Dean, A.; Kim, A. The role of transcriptional activator GATA-1 at human  $\gamma$ -globin HS2. *Nucleic Acids Res.* **2008**, *36*, 4521–4528. [[CrossRef](#)] [[PubMed](#)]
36. Im, H.; Grass, J.A.; Johnson, K.D.; Kim, S.-I.; Boyer, M.E.; Imbalzano, A.N.; Bieker, J.J.; Bresnick, E.H. Chromatin domain activation via GATA-1 utilization of a small subset of dispersed GATA motifs within a broad chromosomal region. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 17065–17070. [[CrossRef](#)] [[PubMed](#)]
37. Sawado, T.; Igarashi, K.; Groudine, M. Activation of beta-major globin gene transcription is associated with recruitment of NF-E2 to the beta-globin LCR and gene promoter. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 10226–10231. [[CrossRef](#)]
38. Gurumurthy, A.; Yu, D.T.; Stees, J.R.; Chamales, P.; Gavrilova, E.; Wassel, P.; Li, L.; Stribling, D.; Chen, J.; Brackett, M.; et al. Super-enhancer mediated regulation of adult  $\beta$ -globin gene expression: The role of eRNA and Integrator. *Nucleic Acids Res.* **2021**, *49*, 1383–1396. [[CrossRef](#)] [[PubMed](#)]
39. Sabari, B.R.; Dall’Agnese, A.; Boija, A.; Klein, I.A.; Coffey, E.L.; Shrinivas, K.; Abraham, B.J.; Hannett, N.M.; Zamudio, A.V.; Manteiga, J.C.; et al. Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **2018**, *361*, eaar3958. [[CrossRef](#)] [[PubMed](#)]
40. Razin, S.V.; Gavrilov, A.A. The Role of Liquid–Liquid Phase Separation in the Compartmentalization of Cell Nucleus and Spatial Genome Organization. *Biochemistry* **2020**, *85*, 643–650. [[CrossRef](#)] [[PubMed](#)]
41. Hewitt, K.; Johnson, K.; Gao, X.; Keles, S.; Bresnick, E. The Hematopoietic Stem and Progenitor Cell Cistrome. *Curr. Top. Dev. Biol.* **2016**, *118*, 45–76. [[CrossRef](#)] [[PubMed](#)]
42. Craddock, C.; Vyas, P.; Sharpe, J.; Ayyub, H.; Wood, W.; Higgs, D. Contrasting effects of alpha and beta globin regulatory elements on chromatin structure may be related to their different chromosomal environments. *EMBO J.* **1995**, *14*, 1718–1726. [[CrossRef](#)] [[PubMed](#)]
43. Forsberg, E.C.; Bresnick, E.H. Histone acetylation beyond promoters: Long-range acetylation patterns in the chromatin world. *BioEssays* **2001**, *23*, 820–830. [[CrossRef](#)] [[PubMed](#)]
44. Forsberg, E.C.; Downs, K.M.; Christensen, H.M.; Im, H.; Nuzzi, P.A.; Bresnick, E.H. Developmentally dynamic histone acetylation pattern of a tissue-specific chromatin domain. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 14494–14499. [[CrossRef](#)]
45. Forrester, W.C.; Epner, E.; Driscoll, M.C.; Enver, T.; Brice, M.; Papayannopoulou, T.; Groudine, M. A deletion of the human b-globin locus activation region causes a major alteration in chromatin structure and replication across the entire b-globin locus. *Gene Dev.* **1990**, *4*, 1637–1649. [[CrossRef](#)] [[PubMed](#)]
46. Vyas, P.; Vickers, M.; Simmons, D.; Ayyub, H.; Craddock, C.; Higgs, D. Cis-acting sequences regulating expression of the human  $\alpha$ -globin cluster lie within constitutively open chromatin. *Cell* **1992**, *69*, 781–793. [[CrossRef](#)]
47. Zhou, G.-L.; Xin, L.; Song, W.; Di, L.-J.; Liu, G.; Wu, X.-S.; Liu, D.-P.; Liang, C.-C. Active Chromatin Hub of the Mouse  $\alpha$ -Globin Locus Forms in a Transcription Factory of Clustered Housekeeping Genes. *Mol. Cell. Biol.* **2006**, *26*, 5096–5105. [[CrossRef](#)]
48. Lee, W.S.; McColl, B.; Maksimovic, J.; Vadolas, J. Epigenetic interplay at the  $\beta$ -globin locus. *Biochim. Biophys. Acta Bioenerg.* **2017**, *1860*, 393–404. [[CrossRef](#)]
49. Fromm, G.; Bulger, M. A spectrum of gene regulatory phenomena at mammalian  $\beta$ -globin gene loci This paper is one of a selection of papers published in this Special Issue, entitled 30th Annual International Asilomar Chromatin and Chromosomes Conference, and has undergone the Journal’s usual peer review process. *Biochem. Cell Biol.* **2009**, *87*, 781–790. [[CrossRef](#)]
50. Kim, Y.W.; Yun, W.J.; Kim, A. Erythroid activator NF-E2, TAL1 and KLF1 play roles in forming the LCR HSs in the human adult  $\beta$ -globin locus. *Int. J. Biochem. Cell Biol.* **2016**, *75*, 45–52. [[CrossRef](#)]
51. Su, M.Y.; Steiner, L.A.; Bogardus, H.; Mishra, T.; Schulz, V.; Hardison, R.C.; Gallagher, P.G. Identification of Biologically Relevant Enhancers in Human Erythroid Cells\*. *J. Biol. Chem.* **2013**, *288*, 8433–8444. [[CrossRef](#)] [[PubMed](#)]
52. Qi, Q.; Cheng, L.; Tang, X.; He, Y.; Li, Y.; Yee, T.; Shrestha, D.; Feng, R.; Xu, P.; Zhou, X.; et al. Dynamic CTCF binding directly mediates interactions among cis-regulatory elements essential for hematopoiesis. *Blood* **2021**, *137*, 1327–1339. [[CrossRef](#)] [[PubMed](#)]
53. Sawaengdee, W.; Cui, K.; Zhao, K.; Hongeng, S.; Fucharoen, S.; Wongtrakoongate, P. Genome-Wide Transcriptional Regulation of the Long Non-coding RNA Steroid Receptor RNA Activator in Human Erythroblasts. *Front. Genet.* **2020**, *11*, 850. [[CrossRef](#)] [[PubMed](#)]
54. Schoenfelder, S.; Sexton, T.; Chakalova, L.; Cope, N.F.; Horton, A.; Andrews, S.; Kurukuti, S.; A Mitchell, J.; Umlauf, D.; Dimitrova, D.S.; et al. Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. *Nat. Genet.* **2009**, *42*, 53–61. [[CrossRef](#)] [[PubMed](#)]
55. Brown, J.M.; Leach, J.; Reittie, J.E.; Atzberger, A.; E Leeprudhoe, J.; Wood, W.G.; Higgs, D.R.; Iborra, F.J.; Buckle, V.J. Coregulated human globin genes are frequently in spatial proximity when active. *J. Cell Biol.* **2006**, *172*, 177–187. [[CrossRef](#)] [[PubMed](#)]
56. Gavrilov, A.A.; Razin, S.V. Spatial configuration of the chicken  $\alpha$ -globin gene domain: Immature and active chromatin hubs. *Nucleic Acids Res.* **2008**, *36*, 4629–4640. [[CrossRef](#)] [[PubMed](#)]
57. Ulianov, S.V.; A Gavrilov, A.; Razin, S.V. Spatial organization of the chicken beta-globin gene domain in erythroid cells of embryonic and adult lineages. *Epigenetics Chromatin* **2012**, *5*, 16. [[CrossRef](#)]
58. Cui, S.; Kolodziej, K.E.; Obara, N.; Amaral-Psarris, A.; Demmers, J.; Shi, L.; Engel, J.D.; Grosveld, F.; Strouboulis, J.; Tanabe, O. Nuclear Receptors TR2 and TR4 Recruit Multiple Epigenetic Transcriptional Corepressors That Associate Specifically with the Embryonic  $\beta$ -Type Globin Promoters in Differentiated Adult Erythroid Cells. *Mol. Cell. Biol.* **2011**, *31*, 3298–3311. [[CrossRef](#)]

59. Guo, X.; Plank-Bazinet, J.; Krivega, I.; Dale, R.K.; Dean, A. Embryonic erythropoiesis and hemoglobin switching require transcriptional repressor ETO2 to modulate chromatin organization. *Nucleic Acids Res.* **2020**, *48*, 10226–10240. [[CrossRef](#)]
60. Liu, N.; Hargreaves, V.V.; Zhu, Q.; Kurland, J.V.; Hong, J.; Kim, W.; Sher, F.; Trevino, C.M.; Rogers, J.M.; Kurita, R.; et al. Direct Promoter Repression by BCL11A Controls the Fetal to Adult Hemoglobin Switch. *Cell* **2018**, *173*, 430–442.e17. [[CrossRef](#)]
61. Yu, L.; Jearawiriyapaisarn, N.; Lee, M.P.; Hosoya, T.; Wu, Q.; Myers, G.; Lim, K.-C.; Kurita, R.; Nakamura, Y.; Vojtek, A.B.; et al. BAP1 regulation of the key adaptor protein NCoR1 is critical for  $\gamma$ -globin gene repression. *Genes Dev.* **2018**, *32*, 1537–1549. [[CrossRef](#)]
62. Iarovaia, O.V.; Kovina, A.P.; Petrova, N.; Razin, S.V.; Ioudinkova, E.S.; Vassetzky, Y.; Ulianov, S.V. Genetic and Epigenetic Mechanisms of  $\beta$ -Globin Gene Switching. *Biochemistry* **2018**, *83*, 381–392. [[CrossRef](#)] [[PubMed](#)]
63. Ulianov, S.V.; Galitsyna, A.A.; Flyamer, I.M.; Golov, A.; Khrameeva, E.E.; Imakaev, M.V.; Abdennur, N.A.; Gelfand, M.S.; Gavrilov, A.A.; Razin, S.V. Activation of the alpha-globin gene expression correlates with dramatic upregulation of nearby non-globin genes and changes in local and large-scale chromatin spatial structure. *Epigenetics Chromatin* **2017**, *10*, 35. [[CrossRef](#)] [[PubMed](#)]
64. Hanssen, L.; Kassouf, M.T.; Oudelaar, A.M.; Biggs, D.; Preece, C.; Downes, D.J.; Gosden, M.; Sharpe, J.A.; Sloane-Stanley, J.A.; Hughes, J.R.; et al. Tissue-specific CTCF-cohesin-mediated chromatin architecture delimits enhancer interactions and function in vivo. *Nat. Cell Biol.* **2017**, *19*, 952–961. [[CrossRef](#)] [[PubMed](#)]
65. Kim, Y.W.; Kang, Y.; Kang, J.; Kim, A. GATA-1-dependent histone H3K27 acetylation mediates erythroid cell-specific chromatin interaction between CTCF sites. *FASEB J.* **2020**, *34*, 14736–14749. [[CrossRef](#)] [[PubMed](#)]
66. Kang, Y.; Kim, Y.W.; Kang, J.; Yun, W.J.; Kim, A. Erythroid specific activator GATA-1-dependent interactions between CTCF sites around the  $\beta$ -globin locus. *Biochim. Biophys. Acta Bioenerg.* **2017**, *1860*, 416–426. [[CrossRef](#)] [[PubMed](#)]
67. Tang, Y.; Wang, Z.; Huang, Y.; Liu, D.-P.; Liu, G.; Shen, W.; Tang, X.; Feng, D.; Liang, C.-C. Gene order in human  $\alpha$ -globin locus is required for their temporal specific expressions. *Genes Cells* **2005**, *11*, 123–131. [[CrossRef](#)] [[PubMed](#)]
68. Harju, S.; Navas, P.A.; Stamatoyannopoulos, G.; Peterson, K.R. Genome Architecture of the Human  $\beta$ -Globin Locus Affects Developmental Regulation of Gene Expression. *Mol. Cell. Biol.* **2005**, *25*, 8765–8778. [[CrossRef](#)] [[PubMed](#)]
69. Tanimoto, K.; Liu, Q.; Bungert, J.; Engel, J.D. Effects of altered gene order or orientation of the locus control region on human  $\beta$ -globin gene expression in mice. *Nature* **1999**, *398*, 344–348. [[CrossRef](#)]
70. A Lloyd, J.; Krakowsky, J.M.; Crable, S.C.; Lingrel, J.B. Human gamma- to beta-globin gene switching using a mini construct in transgenic mice. *Mol. Cell. Biol.* **1992**, *12*, 1561–1567. [[CrossRef](#)]
71. Martin, D.I.; Fiering, S.; Groudine, M. Regulation of  $\beta$ -globin gene expression: Straightening out the locus. *Curr. Opin. Genet. Dev.* **1996**, *6*, 488–495. [[CrossRef](#)]
72. Soshnikova, N.; Duboule, D. Epigenetic Temporal Control of Mouse Hox Genes in Vivo. *Science* **2009**, *324*, 1320–1323. [[CrossRef](#)] [[PubMed](#)]
73. Montavon, T.; Soshnikova, N. Hox gene regulation and timing in embryogenesis. *Semin. Cell Dev. Biol.* **2014**, *34*, 76–84. [[CrossRef](#)] [[PubMed](#)]
74. Duboule, D. The rise and fall of Hox gene clusters. *Development* **2007**, *134*, 2549–2560. [[CrossRef](#)] [[PubMed](#)]
75. Gonzalez, F.; Duboule, D.; Spitz, F. Transgenic analysis of Hoxd gene regulation during digit development. *Dev. Biol.* **2007**, *306*, 847–859. [[CrossRef](#)]
76. Spitz, F.; Gonzalez, F.; Duboule, D. A Global Control Region Defines a Chromosomal Regulatory Landscape Containing the HoxD Cluster. *Cell* **2003**, *113*, 405–417. [[CrossRef](#)]
77. Andrey, G.; Montavon, T.; Mascrez, B.; Gonzalez, F.; Noordermeer, D.; Leleu, M.; Trono, D.; Spitz, F.; Duboule, D. A Switch Between Topological Domains Underlies HoxD Genes Collinearity in Mouse Limbs. *Science* **2013**, *340*, 1234167. [[CrossRef](#)]
78. Fabre, P.J.; Leleu, M.; Mormann, B.H.; Lopez-Delisle, L.; Noordermeer, D.; Beccari, L.; Duboule, D. Large scale genomic reorganization of topological domains at the HoxD locus. *Genome Biol.* **2017**, *18*, 1–15. [[CrossRef](#)]
79. Rodríguez-Carballo, E.; Lopez-Delisle, L.; Willemin, A.; Beccari, L.; Gitto, S.; Mascrez, B.; Duboule, D. Chromatin topology and the timing of enhancer function at the HoxD locus. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 31231–31241. [[CrossRef](#)] [[PubMed](#)]
80. Rodríguez-Carballo, E.; Lopez-Delisle, L.; Yakushiji-Kaminatsui, N.; Ullate-Agote, A.; Duboule, D. Impact of genome architecture on the functional activation and repression of Hox regulatory landscapes. *BMC Biol.* **2019**, *17*, 1–18. [[CrossRef](#)]
81. Khan, M.; Vaes, E.; Mombaerts, P. Regulation of the Probability of Mouse Odorant Receptor Gene Choice. *Cell* **2011**, *147*, 907–921. [[CrossRef](#)]
82. Fuss, S.H.; Omura, M.; Mombaerts, P. Local and cis Effects of the H Element on Expression of Odorant Receptor Genes in Mouse. *Cell* **2007**, *130*, 373–384. [[CrossRef](#)]
83. Serizawa, S.; Miyamichi, K.; Nakatani, H.; Suzuki, M.; Saito, M.; Yoshihara, Y.; Sakano, H. Negative Feedback Regulation Ensures the One Receptor-One Olfactory Neuron Rule in Mouse. *Science* **2003**, *302*, 2088–2094. [[CrossRef](#)]
84. Monahan, K.; Horta, A.; Lomvardas, S. LHX2- and LDB1-mediated trans interactions regulate olfactory receptor choice. *Nat. Cell Biol.* **2019**, *565*, 448–453. [[CrossRef](#)] [[PubMed](#)]
85. Bashkirova, E.; Lomvardas, S. Olfactory receptor genes make the case for inter-chromosomal interactions. *Curr. Opin. Genet. Dev.* **2019**, *55*, 106–113. [[CrossRef](#)] [[PubMed](#)]
86. Iwata, T.; Niimura, Y.; Kobayashi, C.; Shirakawa, D.; Suzuki, H.; Enomoto, T.; Touhara, K.; Yoshihara, Y.; Hirota, J. A long-range cis-regulatory element for class I odorant receptor genes. *Nat. Commun.* **2017**, *8*, 885. [[CrossRef](#)] [[PubMed](#)]

87. Tsai, Y.-C.; Cooke, N.E.; Liebhaber, S.A. Long-range looping of a locus control region drives tissue-specific chromatin packing within a multigene cluster. *Nucleic Acids Res.* **2016**, *44*, 4651–4664. [[CrossRef](#)] [[PubMed](#)]
88. Jia, Z.; Li, J.; Ge, X.; Wu, Y.; Guo, Y.; Wu, Q. Tandem CTCF sites function as insulators to balance spatial chromatin contacts and topological enhancer-promoter selection. *Genome Biol.* **2020**, *21*, 1–24. [[CrossRef](#)]
89. Guo, Y.; Monahan, K.; Wu, H.; Gertz, J.; Varley, K.; Li, W.; Myers, R.M.; Maniatis, T.; Wu, Q. CTCF/cohesin-mediated DNA looping is required for protocadherin promoter choice. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 21081–21086. [[CrossRef](#)]
90. Canzio, D.; Nwakeze, C.L.; Horta, A.; Rajkumar, S.M.; Coffey, E.L.; Duffy, E.E.; Duffié, R.; Monahan, K.; O’Keeffe, S.; Simon, M.D.; et al. Antisense lncRNA Transcription Mediates DNA Demethylation to Drive Stochastic Protocadherin  $\alpha$  Promoter Choice. *Cell* **2019**, *177*, 639–653. [[CrossRef](#)]
91. Almenar-Queralt, A.; Merkurjev, D.; Kim, H.S.; Navarro, M.; Ma, Q.; Chaves, R.S.; Allegue, C.; Driscoll, S.P.; Chen, A.G.; Kohlnhofer, B.; et al. Chromatin establishes an immature version of neuronal protocadherin selection during the naive-to-primed conversion of pluripotent stem cells. *Nat. Genet.* **2019**, *51*, 1691–1701. [[CrossRef](#)] [[PubMed](#)]
92. Zhang, T.; Haws, P.; Wu, Q. Multiple Variable First Exons: A Mechanism for Cell- and Tissue-Specific Gene Regulation. *Genome Res.* **2003**, *14*, 79–89. [[CrossRef](#)] [[PubMed](#)]
93. Hozumi, N.; Tonegawa, S. Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. 1976 [classical article]. *J. Immunol.* **2004**, *173*, 4260–4264. [[PubMed](#)]
94. Michalak, P. Coexpression, coregulation, and cofunctionality of neighboring genes in eukaryotic genomes. *Genomics* **2008**, *91*, 243–248. [[CrossRef](#)] [[PubMed](#)]
95. Purmann, A.; Toedling, J.; Schueler, M.; Carninci, P.; Lehrach, H.; Hayashizaki, Y.; Huber, W.; Sperling, S. Genomic organization of transcriptomes in mammals: Coregulation and cofunctionality. *Genomics* **2007**, *89*, 580–587. [[CrossRef](#)]
96. Sémon, M.; Duret, L. Evolutionary Origin and Maintenance of Coexpressed Gene Clusters in Mammals. *Mol. Biol. Evol.* **2006**, *23*, 1715–1723. [[CrossRef](#)]
97. Le Dily, F.; Baù, D.; Pohl, A.; Vicent, G.P.; Serra, F.; Soronellas, D.; Castellano, G.; Wright, R.H.; Ballare, C.; Fillion, G.; et al. Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. *Genes Dev.* **2014**, *28*, 2151–2162. [[CrossRef](#)]
98. Le Dily, F.; Beato, M. TADs as modular and dynamic units for gene regulation by hormones. *FEBS Lett.* **2015**, *589*, 2885–2892. [[CrossRef](#)]
99. Sun, F.; Chronis, C.; Kronenberg, M.; Chen, X.-F.; Su, T.; Lay, F.D.; Plath, K.; Kurdistani, S.K.; Carey, M.F. Promoter-Enhancer Communication Occurs Primarily within Insulated Neighborhoods. *Mol. Cell* **2019**, *73*, 250–263.e5. [[CrossRef](#)]
100. Soler-Oliva, M.E.; Guerrero-Martínez, J.A.; Bachetti, V.; Reyes, J.C. Analysis of the relationship between coexpression domains and chromatin 3D organization. *PLoS Comput. Biol.* **2017**, *13*, e1005708. [[CrossRef](#)]
101. Storz, J.F.; Opazo, J.C.; Hoffmann, F.G. Gene duplication, genome duplication, and the functional diversification of vertebrate globins. *Mol. Phylogenetics Evol.* **2013**, *66*, 469–478. [[CrossRef](#)] [[PubMed](#)]
102. Hardison, R.C. Globin genes on the move. *J. Biol.* **2008**, *7*, 35. [[CrossRef](#)]
103. Hardison, R.C. A brief history of hemoglobins: Plant, animal, protist, and bacteria. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 5675–5679. [[CrossRef](#)] [[PubMed](#)]
104. Blank, M.; Wollberg, J.; Gerlach, F.; Reimann, K.; Roesner, A.; Hankeln, T.; Fago, A.; Weber, R.E.; Burmester, T. A Membrane-Bound Vertebrate Globin. *PLoS ONE* **2011**, *6*, e25292. [[CrossRef](#)] [[PubMed](#)]
105. Ohno, S. *Evolution by Gene Duplication*; Springer: Berlin/Heidelberg, Germany, 1970; ISBN 978-3-642-86661-6.
106. Burmester, T.; Hankeln, T. Function and evolution of vertebrate globins. *Acta Physiol.* **2014**, *211*, 501–514. [[CrossRef](#)]
107. Schwarze, K.; Burmester, T. Conservation of globin genes in the “living fossil” *Latimeria chalumnae* and reconstruction of the evolution of the vertebrate globin family. *Biochim. Biophys. Acta Proteins Proteom.* **2013**, *1834*, 1801–1812. [[CrossRef](#)]
108. Lechauve, C.; Jager, M.; Laguerre, L.; Kiger, L.; Correc, G.; Leroux, C.; Vinogradov, S.; Czjzek, M.; Marden, M.C.; Bailly, X. Neuroglobins, Pivotal Proteins Associated with Emerging Neural Systems and Precursors of Metazoan Globin Diversity. *J. Biol. Chem.* **2013**, *288*, 6957–6967. [[CrossRef](#)]
109. Braasch, I.; Brunet, F.; Volff, J.-N.; Scharl, M. Pigmentation Pathway Evolution after Whole-Genome Duplication in Fish. *Genome Biol. Evol.* **2009**, *1*, 479–493. [[CrossRef](#)] [[PubMed](#)]
110. Braasch, I.; Volff, J.-N.; Scharl, M. The Endothelin System: Evolution of Vertebrate-Specific Ligand-Receptor Interactions by Three Rounds of Genome Duplication. *Mol. Biol. Evol.* **2009**, *26*, 783–799. [[CrossRef](#)]
111. Holland, P.; Garcia-Fernández, J.; Williams, N.A.; Sidow, A. Gene duplications and the origins of vertebrate development. *Development.* **1994**, *1994*, 125–133. [[CrossRef](#)]
112. Meyer, A. Hox gene variation and evolution. *Nat. Cell Biol.* **1998**, *391*, 227–228. [[CrossRef](#)]
113. Meyer, A.; Scharl, M. Gene and genome duplications in vertebrates: The one-to-four (to-eight in fish) rule and the evolution of novel gene functions. *Curr. Opin. Cell Biol.* **1999**, *11*, 699–704. [[CrossRef](#)]
114. Shimeld, S.M.; Holland, P. Vertebrate innovations. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 4449–4452. [[CrossRef](#)] [[PubMed](#)]
115. Wada, H.; Makabe, K. Genome duplications of early vertebrates as a possible chronicle of the evolutionary history of the neural crest. *Int. J. Biol. Sci.* **2006**, *2*, 133–141. [[CrossRef](#)] [[PubMed](#)]
116. Zhang, G.; Cohn, M.J. Genome duplication and the origin of the vertebrate skeleton. *Curr. Opin. Genet. Dev.* **2008**, *18*, 387–393. [[CrossRef](#)] [[PubMed](#)]

117. Hoffmann, F.G.; Opazo, J.C.; Storz, J.F. Whole-Genome Duplications Spurred the Functional Diversification of the Globin Gene Superfamily in Vertebrates. *Mol. Biol. Evol.* **2011**, *29*, 303–312. [[CrossRef](#)] [[PubMed](#)]
118. Storz, J.F. Gene Duplication and Evolutionary Innovations in Hemoglobin-Oxygen Transport. *Physiology* **2016**, *31*, 223–232. [[CrossRef](#)] [[PubMed](#)]
119. Goodman, M.; Moore, G.W.; Barnabas, J.; Matsuda, G. The phylogeny of human globin genes investigated by the maximum parsimony method. *J. Mol. Evol.* **1974**, *3*, 1–48. [[CrossRef](#)]
120. Jeffreys, A.J.; Wilson, V.; Wood, D.; Simons, J.P.; Kay, R.M.; Williams, J.G. Linkage of adult alpha- and beta-globin genes in *X. laevis* and gene duplication by tetraploidization. *Cell* **1980**, *21*, 555–564. [[CrossRef](#)]
121. Hosbach, H.A.; Wylter, T.; Weber, R. The xenopus laevis globin gene family: Chromosomal arrangement and gene structure. *Cell* **1983**, *32*, 45–53. [[CrossRef](#)]
122. Patient, R.K.; Elkington, J.A.; Kay, R.M.; Williams, J.G. Internal organization of the major adult  $\alpha$ - and  $\beta$ -globin genes of *X. laevis*. *Cell* **1980**, *21*, 565–573. [[CrossRef](#)]
123. Queiroz, J.P.F.; Lima, N.C.B.; Rocha, B.A.M. The rise and fall of globins in the amphibia. *Comp. Biochem. Physiol. Part D Genom. Proteom.* **2021**, *37*, 100759. [[CrossRef](#)]
124. Kovina, A.P.; Petrova, N.; Gushchanskaya, E.S.; Dolgushin, K.V.; Gerasimov, E.; Galitsyna, A.A.; Penin, A.A.; Flyamer, I.; Ioudinkova, E.S.; Gavrilov, A.A.; et al. Evolution of the Genome 3D Organization: Comparison of Fused and Segregated Globin Gene Clusters. *Mol. Biol. Evol.* **2017**, *34*, 1492–1504. [[CrossRef](#)] [[PubMed](#)]
125. Patel, V.S.; Cooper, S.J.; E Deakin, J.; Fulton, B.; Graves, T.; Warren, W.C.; Wilson, R.K.; Graves, J.A. Platypus globin genes and flanking loci suggest a new insertional model for beta-globin evolution in birds and mammals. *BMC Biol.* **2008**, *6*, 34. [[CrossRef](#)] [[PubMed](#)]
126. Patel, V.S.; Ezaz, T.; Deakin, J.E.; Graves, J.A.M. Globin gene structure in a reptile supports the transpositional model for amniote  $\alpha$ - and  $\beta$ -globin gene evolution. *Chromosom. Res.* **2010**, *18*, 897–907. [[CrossRef](#)] [[PubMed](#)]
127. Flint, J.; Tufarelli, C.; Peden, J.; Clark, K.; Daniels, R.J.; Hardison, R.; Miller, W.; Philipson, S.; Tan-Un, K.C.; McMorrow, T.; et al. Comparative genome analysis delimits a chromosomal domain and identifies key regulatory elements in the alpha globin cluster. *Hum. Mol. Genet.* **2001**, *10*, 371–382. [[CrossRef](#)]
128. Hoffmann, F.G.; Vandeweghe, M.W.; Storz, J.F.; Opazo, J.C. Gene Turnover and Diversification of the  $\alpha$ - and  $\beta$ -Globin Gene Families in Sauropsid Vertebrates. *Genome Biol. Evol.* **2018**, *10*, 344–358. [[CrossRef](#)]
129. Opazo, J.C.; Hoffmann, F.G.; Storz, J.F. Genomic evidence for independent origins of  $\omega$ -like globin genes in monotremes and therian mammals. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 1590–1595. [[CrossRef](#)] [[PubMed](#)]
130. Wheeler, D.; Hope, R.M.; Cooper, S.J.; Gooley, A.A.; Holland, R.A. Linkage of the beta-like omega-globin gene to alpha-like globin genes in an Australian marsupial supports the chromosome duplication model for separation of globin gene clusters. *J. Mol. Evol.* **2004**, *58*, 642–652. [[CrossRef](#)] [[PubMed](#)]
131. Rodin, S.N.; Parkhomchuk, D.V.; Rodin, A.S.; Holmquist, G.P.; Riggs, A.D. Repositioning-Dependent Fate of Duplicate Genes. *DNA Cell Biol.* **2005**, *24*, 529–542. [[CrossRef](#)] [[PubMed](#)]
132. Philonenko, E.S.; Klochkov, D.B.; Borunova, V.V.; Gavrilov, A.A.; Razin, S.V.; Iarovaia, O.V. TMEM8—A non-globin gene entrapped in the globin web. *Nucleic Acids Res.* **2009**, *37*, 7394–7406. [[CrossRef](#)] [[PubMed](#)]
133. Prioleau, M.-N.; Nony, P.; Simpson, M.; Felsenfeld, G. An insulator element and condensed chromatin region separate the chicken  $\beta$ -globin locus from an independently regulated erythroid-specific folate receptor gene. *EMBO J.* **1999**, *18*, 4035–4048. [[CrossRef](#)] [[PubMed](#)]