Gene Regulation in the Third Dimension

Job Dekker

Science **319**, 1793 (2008);
DOI: 10.1126/science.1152850

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005.

PERSPECTIVE

Gene Regulation in the Third Dimension

Job Dekker*

Analysis of the spatial organization of chromosomes reveals complex three-dimensional networks of chromosomal interactions. These interactions affect gene expression at multiple levels, including long-range control by distant enhancers and repressors, coordinated expression of genes, and modification of epigenetic states. Major challenges now include deciphering the mechanisms by which loci come together and understanding the functional consequences of these often transient associations.

In compact genomes of organisms such as yeast, a gene and its regulatory elements form an uninterrupted genomic segment that constitutes a "regulatory expression unit." However, in more complex genomes, such as those of human and mouse, genes and their regulatory elements can be dispersed over many hundreds of kilobases (*1*, *2*). It has long been hypothesized that communication between widely spaced genomic elements can be facilitated by the spatial organization of chromosomes that brings genes and their regulatory elements in close proximity (Fig. 1A).

The organization of chromosomes has been studied by microscopy and more recently by chromosome conformation capture (3C) (*3*). 3C is a molecular technique that uses formaldehyde cross-linking and locus-specific PCR to detect physical contacts between genomic loci. These approaches are complementary, with microscopy providing information on single cells, but with relatively low resolution, and 3C allowing much higher resolution analyses, but requiring larger cell populations. 3C and microscopy studies confirm that long-range chromosomal interactions are widespread, which suggests a high level of communication between dispersed genomic elements.

Spatial Assembly of Expression Units

Well-characterized examples of spatial association of genomic elements involve interactions between enhancers and target genes. An example is that of the β-globin locus. The locus contains several β-globin–like genes that are regulated by a single cis-acting element, the locus control region (LCR), which is located about 10 to 60 kb upstream of the globin gene promoters. The LCR was found to physically associate with the active globin gene (*4*). Many more examples of long-range looping events have been identified, e.g., in the α-globin locus (*5*, *6*) and the interleukin gene cluster (*7*), and also in controlling single genes [e.g. (*8*, *9*)].

Highly specific associations between loci located on separate chromosomes have also been described. These trans-interactions can be between enhancers and putative target genes, as in the case of olfactory receptor genes (*10*). However, in other cases, they appear to play a role in a higher level of gene control to coordinately regulate multiple loci (Fig. 1B). One example is the association between the T helper 2 cytokine locus on mouse chromosome 11 and the interferon γ gene on chromosome 10 (*11*). Expression of these loci is mutually exclusive, and interaction between them may provide an opportunity to initiate or enforce opposite epigenetic states.

The process of mammalian X-chromosome inactivation involves a specific trans-association between the X chromosomes. Female cells carry two X chromosomes, one of which is mostly silenced so that expression levels of X-linked genes are comparable to those in male cells. X inactivation is initiated at the X-inactivation center. Recently, a transient interaction between the two X-inactivation centers was detected during the developmental stages at which X inactivation is initiated (*12*, *13*). Analysis of mutations in the inactivation centers showed that their association is intimately involved in the X-inactivation process. X chromosome pairing provides an elegant mechanism for counting the number of X chromosomes and for ensuring that their epigenetic fates are linked so that when one chromosome is inactivated the other is not.

These observations suggest an interesting model of what constitutes a "regulatory expression unit" in complex genomes. Whereas in compact genomes, genes and their regulatory elements cluster along the linear genome sequence, in more complex genomes, expression units can be assembled by spatial clustering of genes and distant regulatory elements (Fig. 1). This mode of de novo assembly of expression units could provide additional levels of gene regulation by allowing combinatorial association of genes and sets of regulatory elements. For example, for imprinted loci, maternal and paternal alleles associate with different elements to assemble into distinct expression units (*14*).

Chromosomal Interactions Are Transient

Many of the observations of long-range interactions have been made using 3C and its variations. Performing 3C is relatively simple, but it has proven more complicated to interpret

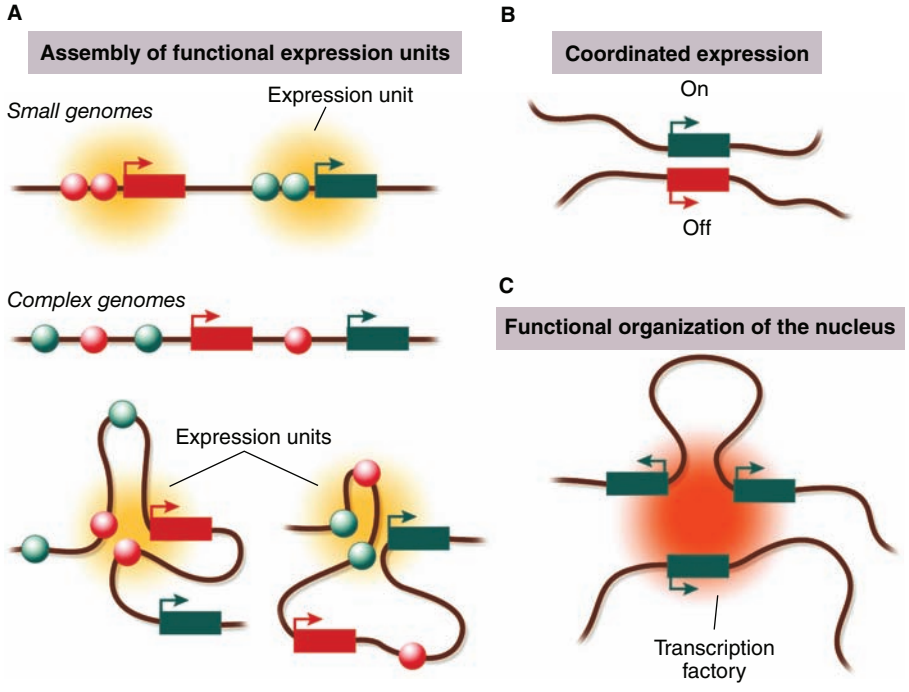

Fig. 1. Spatial assemblies. (**A**) Linearly defined expression units in compact genomes and spatially assembled expression units in complex genomes. (**B**) Association between coordinately expressed genes. (**C**) Colocalization of genes at subnuclear structures, such as transcription factories. Circles, regulatory elements; rectangles, genes. Arrows indicate direction of transcription.

Program in Gene Function and Expression and Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, 364 Plantation Street, Worcester, MA 01605–0103, USA.

*To whom correspondence should be addressed. E-mail: Job.Dekker@umassmed.edu

the results, as has been discussed in several reviews (*15*, *16*). In particular, although many of the chromosomal interactions detected with 3C have been confirmed by microscopy, it is difficult to relate 3C signals to actual frequencies of association. In many cases, the frequency of colocalization is rather low (less than 10% of cells at a given point in time), in accordance with the fact that chromosome conformation is dynamic and highly variable among individual cells. Therefore, the common use of rather rigid looping models to describe these associations, although appealing, can be misleading because these models do not reflect the highly transient nature of long-range interactions.

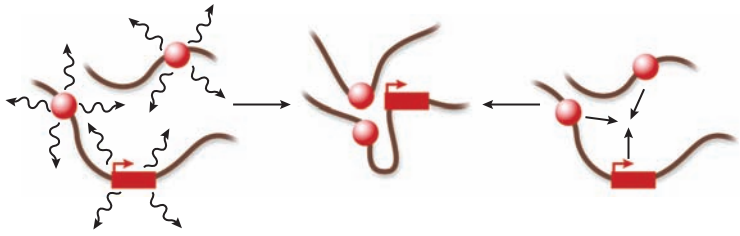

Fig. 2. Passive and active models for bringing loci together. Circles and rectangles represent regulatory elements and genes. Wavy arrows indicate random diffusion. Straight arrows indicate active and directed movement.

Functions of Chromosomal Interactions

Observing a specific association between two loci does not by itself reveal a function for that interaction. Additional approaches such as knock-down of proteins (e.g., transcription factors) that mediate the interaction or deletion of the regulatory element can reveal causal relations between long-range interactions and gene regulation. Another powerful approach is to analyze colocalization of loci by in situ hybridization combined with simultaneous visualization of RNA production at the gene to determine whether the interaction is correlated in time with gene transcription at the level of single cells. It should be noted that interactions have been observed that correlated with gene transcription but that deletion of the interacting regulatory element did not affect expression (*10*, *17*). Although this could indicate that the interaction is not relevant, it could equally reflect our very limited understanding of the role of chromosomal associations in genome regulation.

How do chromosomal associations affect gene expression? Enhancer-promoter interactions could aid in stable recruitment of components of the transcription machinery to the promoter. In addition, enhancer-bound enzymatic activities could be brought in contact with promoter complexes that are then modified, e.g., phosphorylated or methylated, which leads to modulation of promoter activity. Other types of interactions, such as those between the X-inactivation centers, could allow coordinated assembly of two distinct protein complexes on the interacting partners. Alternatively, given the very transient nature of these associations, the two loci may acquire distinct but stable marks, e.g., DNA methylation, that direct assembly of protein complexes at later time points when the loci no longer interact.

How Do Loci Get Together?

Several models have been proposed (*18*) by which distant genomic elements contact each other (Fig. 2). Passive diffusion models are based on the assumption that the mobility of loci provides opportunities for random collisions that are then converted into productive interactions; whether they are productive is dependent on the affinity and specificity of bound protein complexes. Although these models are appealing, it seems that active processes are required, as well, to directly guide loci toward each other. For instance, enhancers have been proposed to actively track along chromatin fibers until a receptive promoter is encountered. Recently, it was found that loci can follow rapid and directed trajectories through the nucleus in an actin-dependent fashion (*19*, *20*). The roles of nuclear actin and myosin have been contentious, but these exciting recent results strongly suggest that they play critical roles in facilitating long-range interactions by transporting loci toward each other or to specific subnuclear neighborhoods, such as transcription factories, which are enriched in RNA polymerase (Fig. 1C).

Genomes also contain regulatory elements that modulate interactions between other loci. So-called insulators prevent an enhancer from activating a promoter but only when it is positioned in between them. How insulators work is not known in detail, but they too engage in long-range interactions with other elements (*21*), which suggests that they generate looped chromosome structures that somehow facilitate the formation of appropriate assemblies of enhancers and target genes.

Future Perspective

At present, significant effort is aimed at comprehensive mapping of chromosomal interactions. Several adaptations of 3C have been developed that allow large-scale detection of genomic interactions by using microarrays or by direct sequencing using any of the newly developed high-throughput sequencing technologies (*22*). The 4C method (3C-on chip, or circular 3C) allows identification of regions throughout the genome that are physically close to a single locus of interest (*23*, *24*). The 5C method (3C–carbon copy) is not anchored on a single locus and is used for mapping dense interaction networks throughout large chromosomal regions of interest (*25*). These approaches will yield new insights into the spatial organization of genomes but are descriptive in nature. Additional approaches will be essential to unravel the mechanisms by which chromosomal associations affect genome regulation. These approaches include time-resolved imaging of chromosomal loci, molecular and genetic manipulation of the mechanisms that control the subnuclear localization and movement of loci, as well as biochemical studies to characterize the complexes that mediate chromosomal associations. Combined, these various approaches promise to provide exciting new insights into the three-dimensional aspects of gene regulation.

References and Notes

1. D. A. Kleinjan, V. van Heyningen, *Am. J. Hum. Genet.* **76**, 8 (2005).
2. A. G. West, P. Fraser, *Hum. Mol. Genet.* **14**, R101 (2005).
3. J. Dekker, K. Rippe, M. Dekker, N. Kleckner, *Science* **295**, 1306 (2002).
4. B. Tolhuis, R. J. Palstra, E. Splinter, F. Grosveld, W. de Laat, *Mol. Cell* **10**, 1453 (2002).
5. D. Vernimmen, M. De Gobbi, J. A. Sloane-Stanley, W. G. Wood, D. R. Higgs, *EMBO J.* **26**, 2041 (2007).
6. G. L. Zhou *et al.*, *Mol. Cell. Biol.* **26**, 5096 (2006).
7. C. G. Spilianakis, R. A. Flavell, *Nat. Immunol.* **5**, 1017 (2004).
8. J. A. Grass *et al.*, *Mol. Cell. Biol.* **26**, 7056 (2006).
9. H. Jing *et al.*, *Mol. Cell* **29**, 232 (2008).
10. S. Lomvardas *et al.*, *Cell* **126**, 403 (2006).
11. C. G. Spilianakis, M. D. Lalioti, T. Town, G. R. Lee, R. A. Flavell, *Nature* **435**, 637 (2005).
12. N. Xu, C. L. Tsai, J. T. Lee, *Science* **311**, 1149 (2006).
13. C. P. Bacher *et al.*, *Nat. Cell Biol.* **8**, 293 (2006).
14. S. Kurukuti *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 10684 (2006).
15. J. Dekker, *Nat. Methods* **3**, 17 (2006).
16. M. Simonis, J. Kooren, W. de Laat, *Nat. Methods* **4**, 895 (2007).
17. S. H. Fuss, M. Omura, P. Mombaerts, *Cell* **130**, 373 (2007).
18. J. D. Engel, K. Tanimoto, *Cell* **100**, 499 (2000).
19. C. H. Chuang *et al.*, *Curr. Biol.* **16**, 825 (2006).
20. M. Dundr *et al.*, *J. Cell Biol.* **179**, 1095 (2007).
21. J. A. Wallace, G. Felsenfeld, *Curr. Opin. Genet. Dev.* **17**, 400 (2007).
22. B. Wold, R. M. Meyer, *Nat. Methods* **5**, 19 (2008).
23. M. Simonis *et al.*, *Nat. Genet.* **38**, 1348 (2006).
24. Z. Zhao *et al.*, *Nat. Genet.* **38**, 1341 (2006).
25. J. Dostie *et al.*, *Genome Res.* **16**, 1299 (2006).
26. J.D. is supported by grants from NIH (HG003143), the Keck Foundation, and the Cystic Fibrosis Foundation. M. Walhout is acknowledged for suggestions for this article.

10.1126/science.1152850