# Gene Regulation: Piecing Together the Puzzle of Enhancer Evolution

**The sequences of some gene regulatory elements diverge considerably, even between closely related species. A detailed analysis of the fast-evolving *sparkling* enhancer in *Drosophila* now identifies key compensatory mechanisms and 'grammar' elements that are critical for maintaining functional integrity.**

**Rhea R. Datta and Stephen Small***

The complex body plans of multicellular organisms are established by networks of genes that establish time- and position-specific patterns of gene expression. Crucial elements in transcriptional networks are enhancers — DNA sequences that control when and where specific genes are expressed. Enhancers contain binding sites for transcription factors, and integrate the activities of all factors that are bound at each time in development [1]. Over time, enhancers produce expression patterns that presage the organization of the mature body plan. Sequence changes affecting enhancer activity are predicted to affect development, and it has been suggested that changes in enhancer function might alter body plans during the course of evolution [2]. However, it is not clear how much sequence change can be tolerated when the activity of the enhancer might affect fitness. One way to address this question is to compare sequences that provide the same functions in different species. A study by Scott Barolo and colleagues [3] in this issue of *Current Biology* looks at an enhancer in *Drosophila* and shows how it generates a robust pattern despite rapid evolution.

Swanson *et al.* [3] analyzed the sequence divergence for seventeen well-characterized enhancers among twelve *Drosophila* species, and found a wide spectrum of variation in the percentage of alignable sequences. Among the most variable enhancers was the 'sparkling' (*spa*) element, which normally drives expression of the Pax2 gene in a subset of cone cells in the developing *Drosophila* eye. This enhancer responds to transcriptional input from the Notch and EGF pathways, and contains 11 characterized binding sites for known factors. Previously, Barolo and coworkers [4] had shown that at least four other regions, outside and in between known sites, are critical for the proper activity of this enhancer, i.e. for the generation of the proper expression pattern in the eye.

Amazingly, the sequence comparison now revealed that only two of the 11 known sites can be easily identified in all twelve species, suggesting a high rate of site turnover. Nonetheless, function is preserved, based on a reporter gene carrying the *spa* element from the distantly related species *Drosophila pseudoobscura* (*Dpse*), which drives faithful cone-cell specific expression in *Drosophila melanogaster* (*Dmel*). Such functional enhancer conservation with sequence divergence had been observed before for yolk protein regulation [5].

## The Pieces of the Puzzle

The authors embarked on a search for the mechanistic details underlying the functional conservation between the *spa* enhancers in *D. pseudoobscura* and *D. melanogaster*, using two hypotheses. First, because the two enhancers generate very similar expression patterns, they might respond to the same set of input factors, some of which may still be unknown. Second, because very few known functional sites from the *Dmel* element are detectable in the *Dpse* sequence, other sequence motifs critical for function might be overrepresented in that sequence.

To find such motifs, they tested the activities of chimeric elements that consisted of the 5′ half of the *Dpse* enhancer fused to the 3′ half of the *Dmel* enhancer, and vice versa. Neither chimera generated the endogenous expression pattern in cone cells, indicating that there is a constrained binding-site architecture that is unique to each species. One chimera (*Dpse5′+Dmel3′*), however, resulted in hyperactivity, which indicated the presence of too many activator sites. As the *Dpse 5′*-half has no enhancer activity on its own, a series of mutations in the *Dmel 3′*-half were tested to identify regions required for augmenting the activation of the chimera. Interestingly, mutating regions in between known regulatory sites in the *Dmel* enhancer prevented activation. Within these regions, the authors identified several sequence motifs that were also present in the *Dpse* sequence, but at different positions. One motif (ε) was particularly interesting because it is present only once in the *Dmel* sequence, but there are nine more or less exact copies in the *Dpse* sequence. Mutations of the single ε-motif in the *Dmel* sequence severely reduced expression. Unfortunately, the identity of the factor that binds the ε site is still unknown.

These data suggest a clear model for how the *Dpse* enhancer can function in the absence of many sites required for *Dmel* enhancer activity. Through its multiple ε sites, the *Dpse* sequence is able to compensate for the loss of sites that are crucial in *D. melanogaster*. The fact that critical binding sites can be in different positions suggests a high level of flexibility in enhancer architecture, and is consistent with the 'billboard' model of enhancer structure, which proposes that enhancers simply summarize binding sites [6].

However, it would be a mistake to classify the *spa* enhancer as a simple billboard that 'counts' binding sites of different types. When site arrangements are compared in the *spa* sequences of different species, small 'grammar elements' — site combinations with spacing constraints — appear to be evolutionarily conserved. These may anchor proteins and facilitate cooperative binding of different transcription factors to the linked sites. Similar grammar elements have been shown to be critical for fine-tuning enhancer activities in the early *Drosophila* embryo [7], and are reminiscent of the strict site arrangements required for 'enhanceosome' function in *interferon-γ* regulation [8].

## Solving the Puzzle in Different Ways

The results of Swanson *et al.* [3] explain nicely how the *spa* enhancer can evolve rapidly without substantially changing its regulatory function. However, it is still unclear why some enhancers change rapidly during evolution while

others remain fairly constant. One possibility is that sequence variation is linked to the type of enhancer grammar required for function.

Enhancers do many different things: they respond to different transacting factor concentrations, integrate different numbers of signals and are active at different times in the development of an animal. Some are amazingly simple, containing tightly-linked sites for as few as two transcription factors, as exemplified by the elements that specify neuronal subtypes in *C. elegans* [9]. In this case, one critical grammar element is sufficient for enhancer activity, and conservation of the basic regulatory mechanism is easy to see.

Other enhancers, specifically those that respond to multiple inputs or different levels of inputs, must require multiple grammar elements, each of which has a precise sub-function. In this type of enhancer, individual elements may be required, but are not sufficient for enhancer function. For example, the well-characterized *even-skipped* stripe 2 enhancer contains at least five grammar elements [10], each of which is critical for its function. However, changes in spacing between elements during evolution have not interfered dramatically with its function [11]. Finally, if specific grammar elements can substitute for each other, then even very different sequences can mediate similar regulatory functions. Since the *spa* enhancer can apparently substitute ε sites for binding sites for the transcriptional regulators Su(H) and Lz [3], it is an excellent example of this most flexible type.

What we know now is that we need to examine enhancer sequences even more closely to identify the as yet elusive motifs *de novo* that can tell us more about regulatory evolution and function. An integrated approach might consist of comparing regulatory elements of coexpressed genes in order to get a glimpse of convergent mechanisms, while studying elements with similar functions across species to see the limits of plasticity in sequence divergence.

The ease of transgenesis and genetic manipulation in *Drosophila* [12] has facilitated efforts toward understanding transcriptional regulation. Detailed analyses like the one of Swanson *et al.* [3] will further enable us to understand the complex language that translates transcriptional inputs into patterns. By combining the vast amount of sequence information available with elegant molecular manipulations, we will be able to take a closer look at the molecular mechanisms that create the astonishing morphological and developmental diversity in multicellular organisms.

**References**

1. Davidson, E.H. (2006). The Regulatory Genome (Oxford, UK: Elsevier).
2. Carroll, S.B. (2008). Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. Cell *134*, 25–36.
3. Swanson, C.I., Schwimmer, D.B., and Barolo, S. (2011). Rapid evolutionary rewiring of a structurally constrained eye enhancer. Curr. Biol. *21*, 1186–1196.
4. Swanson, C.I., Evans, N.C., and Barolo, S. (2010). Structural rules and complex regulatory circuitry constrain expression of a Notch- and EGFR-regulated eye enhancer. Dev. Cell *18*, 359–370.
5. Piano, F., Parisi, M.J., Karess, R., and Kambysellis, M.P. (1999). Evidence for redundancy but not trans factor-cis element coevolution in the regulation of Drosophila Yp genes. Genetics *152*, 605–616.
6. Arnosti, D.N. (2003). Analysis and function of transcriptional regulatory elements: insights from Drosophila. Annu. Rev. Entomol. *48*, 579–602.
7. Zinzen, R.P., Senger, K., Levine, M., and Papatsenko, D. (2006). Computational models for neurogenic gene expression in the Drosophila embryo. Curr. Biol. *16*, 1358–1365.
8. Kim, T.K., and Maniatis, T. (1997). The mechanism of transcriptional synergy of an in vitro assembled interferon-beta enhanceosome. Mol. Cell *1*, 119–129.
9. Flames, N., and Hobert, O. (2009). Gene regulatory logic of dopamine neuron differentiation. Nature *458*, 885–889.
10. Arnosti, D.N., Barolo, S., Levine, M., and Small, S. (1996). The eve stripe 2 enhancer employs multiple modes of transcriptional synergy. Development *122*, 205–214.
11. Ludwig, M.Z., Patel, N.H., and Kreitman, M. (1998). Functional analysis of eve stripe 2 enhancer evolution in Drosophila: rules governing conservation and change. Development *125*, 949–958.
12. Venken, K.J., and Bellen, H.J. (2007). Transgenesis upgrades for Drosophila melanogaster. Development *134*, 3571–3584.

Department of Biology, New York University, 100 Washington Square East, New York, NY 10003, USA.
*E-mail: sjs1@nyu.edu

DOI: 10.1016/j.cub.2011.06.026

# Speech Perception: A Language-Trained Chimpanzee Weighs In

**A language-trained chimpanzee is able to interpret synthetic 'auditory caricatures' as speech. Important components of human speech perception thus rely upon general auditory mechanisms that predated the evolution of spoken language.**

**W. Tecumseh Fitch**

It is a commonplace that most dogs recognize their name and a few special words, like 'walk' or 'dinner'. In extraordinary cases, dogs can learn to recognize hundreds of words [1]. It thus comes as no great surprise that a chimpanzee raised in close contact with humans can also recognize hundreds of spoken words, as documented in this issue of *Current Biology* by Heimbauer, Beran and Owren [2] in a study with a common chimpanzee called 'Panzee'.

What has remained unclear, for many years, is whether the same perceptual mechanisms are used in speech recognition by humans and animals. It may be, for example, that a dog recognizes its name simply by intonation pattern, rather than using the detailed phonetic cues we humans rely upon. In other words, if you changed the phonemes of 'Fido' to 'Ginger', but used the same pitch contour, your dog might not even notice the difference. The new study [2] shows that, at least for chimpanzees, the similarities with human speech perception are far deeper and more pervasive than that. This discovery has important implications for a long-running debate in speech science about the evolutionary relationship between production and perception.

Human speech perception and production are like mirrors: our capabilities in the two domains are remarkably well-matched. If I clearly