

ACHIEVEMENTS

Ancient Genetic “Start” Signal Found Across Three Domains of Life

Share:     

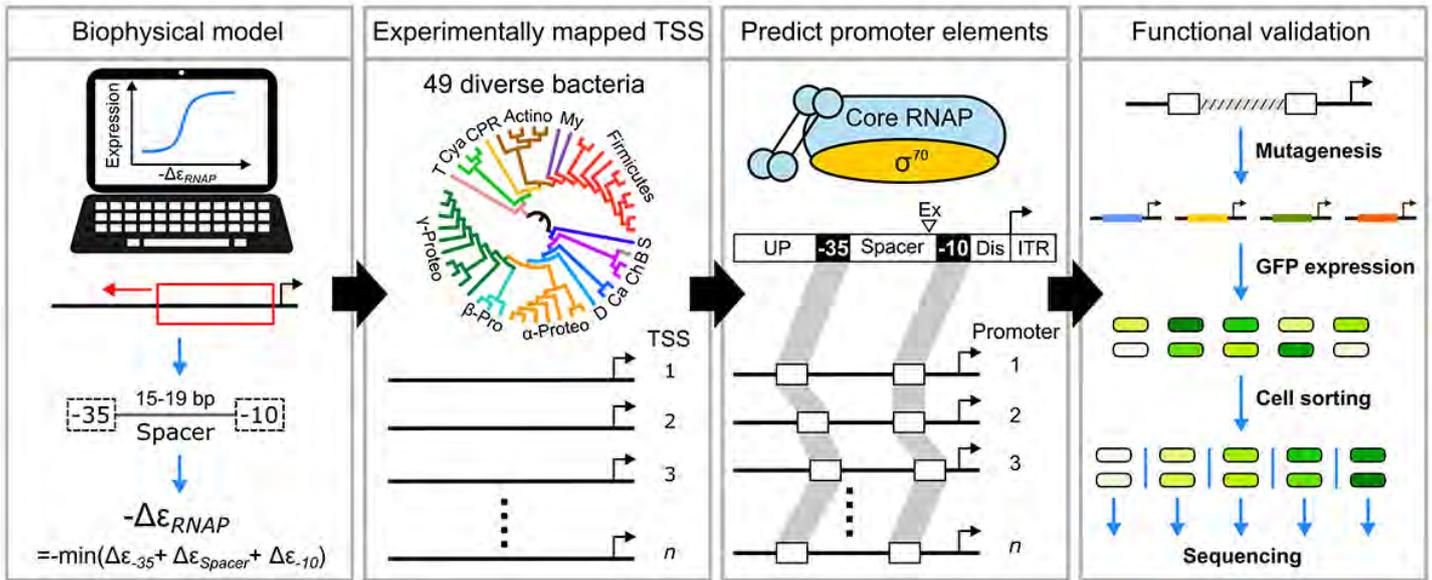


Fig. 1: Schematic of the systems biology approach for decoding bacterial promoters.

DNA may be regarded as the language of life. Using four letters—A, T, G, and C—genomes encode thousands of genes and regulatory elements that govern gene expression. Decoding DNA sequences not only illuminates the origins of life but also helps explain how organisms adapt to changing environments, providing essential knowledge for synthetic biology and for predicting future evolutionary trajectories.

Since genes follow strict rules for encoding proteins, identifying them in DNA is relatively straightforward. Regulatory elements, by contrast, vary widely in length and sequence, making them far more difficult to detect. While tools such as AlphaFold have revolutionized protein structure prediction and design, a major frontier in the life sciences remains the development of computational models capable of predicting and engineering regulatory DNA elements.

Among these regulatory elements, promoters play a central role by controlling transcription—the first step in gene expression. To decipher the sequence rules underlying the promoter function, Prof. David Chou of the Department of Life Science has applied high-throughput approaches to characterize more than 16 million promoter variants in *E. coli* (Fig. 1). This extensive dataset was mined to

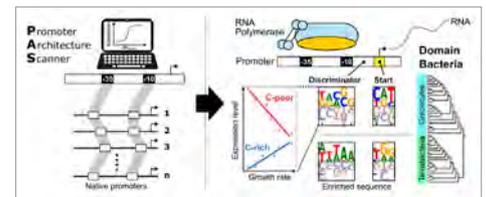


Fig. 3: Cover of *Nucleic Acids Research*, highlighting the discovery of the “start” promoter element conserved across the bacterial domain. Together with the -35 and -10 elements, the start element is recognized by RNA polymerase to initiate transcription “start” series.

train a predictive model, which in turn was applied to analyze 49 diverse bacterial genomes. The study revealed a broadly conserved promoter architecture and yielded two major findings (Fig. 2):

1. In addition to the well-known “-35” and “-10” elements, bacterial promoters contain a previously unrecognized and widely conserved element that specifies transcription start sites. Prof. Chou termed this newly identified element “start”.
2. The “discriminator” element, located downstream of the -10 element, exhibits pronounced sequence divergence between the two major bacterial clades, Terrabacteria and Gracilicutes. This divergence reflects a functional distinction: many Gracilicutes use discriminator sequences to tune promoter strength in response to growth rate, whereas Terrabacteria do not employ this regulatory mechanism.

Remarkably, the bacterial “start” element closely resembles the “initiator” element used by archaea and eukaryotes to define transcription start sites. This similarity suggests that the last universal common ancestor may have relied on a promoter architecture akin to that still observed today, offering the promise of fresh insights into the ancient origins of gene regulation on Earth.

Prof. Chou observed, “These findings reshape our understanding of evolution and lay the foundation for identifying and engineering regulatory elements in microbial genomes.” The study report was selected as the cover story of *Nucleic Acids Research*, Volume 53, Issues 21–22 (Fig. 3).



Fig. 3: Cover of *Nucleic Acids Research*, highlighting the discovery of the “start” promoter element conserved across the bacterial domain. Together with the -35 and -10 elements, the start element is recognized by RNA polymerase to initiate transcription “start” series.



Click or Scan the QR code to read the full article in *Nucleic Acids Research*.