



Professor Ye Kai is a full-time Professor at the Faculty of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, China.

Seq2image: A Holistic Image-Based Paradigm for Genomic Sequence Analysis

Abstract: The analysis of complex biological sequences has historically been dominated by reductionist approaches, wherein elegant statistical models are iteratively patched with numerous ad-hoc filters to handle real-world data complexity, ultimately leading to over-engineered pipelines. To bridge the gap between reductionism and holism, we present seq2image, a conceptual framework that transforms one-dimensional biological sequences into two-dimensional images, thereby leveraging deep learning and advanced image processing techniques to encode both raw data and biological knowledge in a data-driven manner.

This talk introduces three major applications of the seq2image paradigm. First, for structural variant (SV) detection, we demonstrate how clustering local alignment features as image patterns enables the resolution of complex SVs (SVision, Nature Methods 2022), the discovery of de novo and somatic variants through inter-sample comparison (SVision-pro, Nature Biotechnology 2024), and population-scale genotyping of novel and pathogenic alleles from pangenome graphs (Swave, Nature Genetics 2026). Second, for gene annotation, we introduce ANNEVO, which models distal genomic interactions via Hi-C-inspired two-dimensional maps and integrates both vertical and horizontal evolutionary histories through a Mixture-of-Experts architecture, achieving highly accurate ab initio annotation at speeds around 100x faster than conventional evidence-dependent pipelines (Nature Methods 2026). Third, by applying seq2image operators to Hi-C data from 1,025 species, we trace the evolution of three-dimensional genome architectures, revealing that checkerboard compartmentalization correlates with multicellular complexity and that plants and animals have undergone distinct evolutionary trajectories characterized by functional partitioning through architectural

transitions (Cell 2026).

Collectively, seq2image provides a unified, scalable framework for genomic discovery, from local variant classification and distal interaction modeling to the quantification of chromosome-scale genome architecture evolution.

Bio:

Ye held previous posts as an Assistant Professor at The Genome Institute at Washington University in St. Louis, USA, as an Assistant Professor at Leiden University Medical Center in the Netherlands, and as a Postdoctoral Researcher at the European Bioinformatics Institute (EBI) in the United Kingdom under Professor Rolf Apweiler. He earned his PhD (cum laude) from Leiden University in the Netherlands, following his Master's degree from the College of Pharmacy and Bachelor's degree from the College of Life Sciences at Wuhan University.

He has been invited to participate in multiple international genome projects, including The 1000 Genomes Project and The Cancer Genome Atlas (TCGA). In early 2016, Ye returned to China and established an interdisciplinary research group integrating informatics and life sciences at Xi'an Jiaotong University.

Ye's research primarily focuses on the development of computational and artificial intelligence-driven methods for genomic sequence analysis, enabling the detection, characterization, and functional interpretation of structural variations, gene structures, and genome architectures. His work spans two interconnected areas: algorithmic innovation and large-scale genomic data analysis.

Research and project experience:

Algorithm development

- **Traditional modeling:** Pindel for structural variant detection; the MSIsensor suite (MSIsensor, MSIsensor-pro, MSIsensor-RNA) for microsatellite instability profiling.
- **Deep learning:** SVision, SVision-pro, and Swave for complex and somatic structural variant discovery via the "seq2image" paradigm; ANNEVO for ab initio gene annotation through modeling of distal chromosomal interactions.

Large-scale data analysis

- **Plants:** Opium poppy genome project.

- **Human:** Chinese Pangenome Consortium for human population diversity.
- **Pan-species:** Darwin Tree of Life project for broad biodiversity characterization.

Selected 10 publications:

1. Yizhuo Che, Stephen J. Bush, Hui Lin, Mingxuan Li, Xiaofei Yang, Qi Xie, Yuchun Liu, Deyu Meng, **Kai Ye*** [The evolution of high-order genome architecture 1 revealed from 1,000 species.](#) *Cell*, DOI: 10.1016/j.cell.2026.03.042
2. Pengyu Zhang, Tun Xu, Songbo Wang, Xiaofei Yang, Peisen Sun, Peng Jia, Bo Wang, Stephen J. Bush, Zemin Ning., **Kai Ye***. (2026) [Highly accurate ab initio gene annotation with ANNEVO.](#) *Nature Methods*, March 12; 23, 740–748.
3. Songbo Wang, Tun Xu, Pengyu Zhang, **Kai Ye***. (2026) [Population-level structural variant characterization from pangenome graph.](#) *Nature Genetics*, March 10; 58, 664–672.
4. Peisen Sun, Stephen J. Bush, Songbo Wang, Peng Jia, Mingxuan Li, Tun Xu, Pengyu Zhang, Xiaofei Yang, Chengyao Wang, Linfeng Xu, Tingjie Wang, **Kai Ye***. (2025) [STMiner: gene-centric spatial transcriptomics for deciphering tumor tissues.](#) *Cell Genomics*, Feb 12;5(2):100771.
5. Dan Xu, Mengyao Guo, Xin Xu, Gan Luo, Yaxin Liu, Stephen J. Bush, Chengyao Wang, Tun Xu, Wenxin Zeng, Chongbing Liao, Qingxia Wang, Wei Zhao, Wenying Zhao, Yuezhuanan Liu, Shanshan Li, Shuangshuang Zhao, Yaming Jiu, Nathalie Sauvonnet, Wuyuan Lu*, Philippe J. Sansonetti*, **Kai Ye***. (2025) [Shigella infection is facilitated by interaction of human enteric \$\alpha\$ -defensin 5 with colonic epithelial receptor P2Y11.](#) *Nature Microbiology*, 10(2), 509-526
6. Songbo Wang, Jiadong Lin, Peng Jia, Tun Xu, Xiujuan Li, Yuezhuanan Liu, Dan Xu, Stephen J. Bush, Deyu Meng, **Kai Ye***. (2024). [De novo and somatic structural variant discovery with SVision-pro.](#) *Nature Biotechnology*, Mar 22; 43, 181–185.
7. Shenghan Gao, Yanyan Jia, Hongtao Guo, Tun Xu, Bo Wang, Stephen J. Bush, Shijie Wan, Yimeng Zhang, Xiaofei Yang, and **Kai Ye*** (2024). [The centromere landscapes of four karyotypically diverse Papaver species provide insights into chromosome evolution and speciation.](#) *Cell Genomics*, 100626.
8. Yang Gao, Xiaofei Yang, Hao Chen, Xinjiang Tan, Zhaoqing Yang, Lian Deng, Baonan Wang, Shuang Kong, Songyang Li, Yuhang Cui, Chang Lei, Yimin Wang, Yuwen Pan, Sen Ma, Hao Sun, Xiaohan Zhao, Yingbing Shi, Ziyi Yang, Dongdong Wu, Shaoyuan Wu, Xingming Zhao, Binyin Shi, Li Jin, Zhibin Hu, Chinese Pangenome Consortium (CPC)*, Yan Lu*, Jiayou Chu*, **Kai Ye***, Shuhua Xu* (2023). [A pangenome reference of 36 Chinese populations.](#) *Nature*, 619(7968), 112-121.
9. Jiadong Lin, Songbo Wang, Peter A. Audano, Deyu Meng, Jacob I. Flores, Walter Kusters, Xiaofei Yang, Peng Jia, Tobias Marschall, Christine R. Beck, **Kai Ye*** (2022). [SVision: a deep learning approach to resolve complex structural variants.](#) *Nature Methods*, 19(10), 1230-1233.
10. Li Guo, Thilo Winzer, Xiaofei Yang, Yi Li, Zemin Ning, Zhesi He, Roxana Teodor, Ying Lu, Tim A. Bowser, Ian A. Graham*, and **Kai Ye*** (2018). [The opium poppy genome and morphinan production.](#) *Science*, 362(6412), 343-347.