

PxBLAT: An Efficient and Ergonomics Python Binding Library for BLAT

Yangyang Li¹ and Rendong Yang¹✉

¹Department of Urology, Northwestern University Feinberg School of Medicine, Chicago, IL 60611

Efficient manipulation and analysis of genomic data is a crucial task in bioinformatics, often relying on tools like the BLAST-Like Alignment Tool (BLAT). However, working with BLAT directly can be computationally intensive and challenging for users less familiar with its interface. This paper introduces PxBLAT, an ergonomic and efficient Python binding library for BLAT, designed to enhance the user experience and performance of genomic analysis tasks. By providing a Pythonic interface to BLAT, PxBLAT simplifies the usage of BLAT, allowing users to incorporate its functionality directly into their Python-based bioinformatics workflows. Our evaluation demonstrates that PxBLAT not only improves the efficiency of BLAT operations but also offers a more intuitive user interface, thereby reducing the complexity of genomic data analysis tasks. Furthermore, our case studies illustrate how PxBLAT enhances both research and educational applications. With its combination of improved efficiency and usability, PxBLAT represents a significant contribution to bioinformatics programming, with the potential to transform genomic data analysis workflows.

Software Libraries | Sequence Analysis | BLAT
Correspondence: rendong.yang@northwestern.edu

Introduction

The continual advancement of genome sequencing technologies has led to an exponential increase in available genomic data. Tools to analyze and manipulate these data have become critically important in both research and clinical contexts. BLAT (BLAST-Like Alignment Tool) is a key utility in bioinformatics, offering high-speed genomic sequence alignments with significant efficiency. Despite its widespread use and acceptance in the bioinformatics community, interfacing with BLAT can present challenges, particularly when integrating it within a broader Python-based analytical pipeline. This paper introduces PxBLAT, a novel Python binding library designed to streamline and enhance the interaction with BLAT, thereby making it more efficient and ergonomic. This binding library serves as a bridge, bringing the high-performance capabilities of BLAT into the Python environment, which is widely regarded for its readability, simplicity, and extensive library support. By improving the usability of BLAT and enhancing its inte-

gration within Python, PxBLAT opens up new possibilities for efficient genomic analysis.

This study details the design, implementation, and evaluation of PxBLAT. We provide evidence of its performance improvements, demonstrate its ergonomic advantages, and discuss its potential applications in genomic research. Through comparison tests and case studies, we highlight how PxBLAT enhances BLAT utilization, potentially contributing to a more streamlined, efficient, and accessible landscape in bioinformatics. The overarching aim of this work is to contribute to the growing body of tools that empower researchers in handling large genomic datasets, enabling faster, more accurate analysis and facilitating breakthroughs in our understanding of genomic data.

Materials and Methods

This is materials and methods. Pybind11 ([Jakob et al., 2016](#)) blat ([Kent, 2002](#))

Design Philosophy of PxBLAT Implementation

Results

PxBLAT has consistent result with BLAT

Benchmarking Performance

Ergonomics

It is possible to add a one-column Figure like this (Figure 2).

Discussion

This is the discussion section where you wax lyrical about your findings.

Acknowledgements

Conflict of interest

Funding

Data availability

The code is available is GitHub [PxBLAT](#). The benchmarking dataset and code is located in GitHub as well.

Jakob, W., Rhinelander, J., and Moldovan, D. pybind11 — seamless operability between c++11 and python, 2016. <https://github.com/pybind/pybind11>.
Kent, W. J. Blat—the blast-like alignment tool. *Genome research*, 12(4):656–664, 2002.



Figure 1. These are cells.

(A) This is a regular figure with a legend as a caption underneath. Inset: 3X zoom. Scale bar, 10 μ m.



Figure 2. This is a nucleus.

(A) This is a one-column figure with a legend as a caption underneath.

Supplementary Information



Figure S1. This is an endosome.

(A) This is a supplementary figure shown as a two-column image with a legend underneath.