bioRxiv preprint doi: https://doi.org/10.1101/2023.08.02.551686; this version posted October 24, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

PxBLAT: An Ergonomic and Efficient Python Binding Library for BLAT

Yangyang Li[1] **and Rendong Yang**[1,2] ✉

[1]Department of Urology, Northwestern University Feinberg School of Medicine, Chicago, IL 60611; [2]Robert H. Lurie Comprehensive Cancer Center, Northwestern University Feinberg School of Medicine, Chicago, IL 60611

Summary: We introduce PxBLAT, a Python library designed to enhance usability and efficiency in interacting with the BLAST-like alignment tool (BLAT). PxBLAT provides an intuitive application programming interface (API) design, allowing the incorporation of its functionality directly into Python-based bioinformatics workflows. Besides, it integrates seamlessly with Biopython and comes equipped with user-centric features like server readiness checks and port retry mechanisms. PxBLAT removes the necessity for system calls and intermediate files, as well as reducing latency and data conversion overhead. Benchmark tests reveal PxBLAT gains a ~20% performance boost compared to BLAT in the Python environment.

Availability and Implementation: PxBLAT supports Python (version 3.9+), and pre-compiled packages are released via PyPI (`https://pypi.org/project/pxblat/`) and Bioconda (`https://anaconda.org/bioconda/pxblat`). The source code and executables are freely available for academic, nonprofit and personal use. Commercial licensing information is available on the Kent Informatics website (`http://www.kentinformatics.com`). Its documentation is available on ReadTheDocs (`https://pxblat.readthedocs.io/en/latest/`).

Software Libraries | Sequence Analysis | BLAT

Correspondence: ***rendong.yang@northwestern.edu***

Introduction

The rise of Python as a preferred programming language within bioinformatics is widely acknowledged as a result of its user-friendly nature, extensive libraries, and unparalleled versatility (Perkel). A variety of libraries have been crafted to augment Python's interface, thereby amplifying the adaptability and compatibility of bioinformatics tools (Putri et al.; Cock et al.). For instance, Biopython (Cock et al.), a preeminent bioinformatics library, furnishes interfaces to tools like Basic Local Alignment Search Tool (BLAST) (Altschul et al.) and Clustal (Higgins and Sharp).

Besides, the relentless progression of genome sequencing technologies has precipitated a dramatic surge in the availability of genomic data. The necessity for tools to decipher and manipulate these data has become paramount within both the research and clinical domains. BLAT (Kent) is a standout within the bioinformatics landscape and is recognized for its capability to swiftly conduct genome sequence alignments. It offers a faster alternative to BLAST (Altschul et al.) for aligning DNA sequences to the human genome (Kent). Despite its widespread use and endorsement within the bioinformatics community, interfacing with BLAT can present challenges. BLAT, implemented in the C language, solely extends utilities via command-line interfaces (CLIs), rendering its integration into Python-based projects inconvenient. Moreover, BLAT triggers system calls and data conversions that cause significant overhead. Until now, a comprehensive approach to enhancing the usability of BLAT has been absent when incorporated into Python-based projects.

Here, we propose PxBLAT, which empowers users to utilize BLAT programmatically, facilitating its seamless integration into novel algorithms or analytical pipelines. PxBLAT serves as a bridge, bringing the high-performance capabilities of BLAT into the Python environment while maintaining reproducibility. The overarching aim of this work is to fill the observed gap by providing a Python binding library specifically tailored for BLAT, addressing both its efficiency and ergonomic concerns.

Implementation

The design of PxBLAT adheres to the principles of readability and simplicity, aiming for an intuitive user experience that diminishes the learning curve. Striving to curtail complexity while enhancing usability and performance, we extracted the implementation of BLAT from the codebase of the UCSC Genome Browser (UCSC), thus reducing dependencies. We retain the original C codebase and reimplement the utilities of BLAT (V37.1) including *faTwoBit*, *gfServer*, and *gfClient* CLI, using the C++ programming language (Kent). Subsequently, we integrate the current C++ code and develop PxBLAT via Pybind11 (Jakob et al., 2016). The approach facilitated direct interaction with the functions of the BLAT without modifying its original source code, thereby preserving the integrity and performance of the BLAT while broadening its capabilities.

The query result of PxBLAT conforms to the

bioRxiv preprint doi: https://doi.org/10.1101/2023.08.02.551686; this version posted October 24, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

QueryResult class of Biopython (Cock et al.), enabling manipulation of the query result using Biopython features (Listing 1). Furthermore, PxBLAT eliminates the need for intermediate files, allowing all operations to be executed in memory. This eradicates the common obstacle of data conversion into specific formats, enabling users to focus more on the core sequence alignment task. Input and output files have been made optional, offering more flexible choices. The use of system calls, though functional, can induce latency and create performance bottlenecks. PxBLAT reduces the need for system calls, thereby improving efficiency (Table 1). Moreover, PxBLAT allows for the retrieval of server status without manipulating log files, a process that can be troublesome in concurrent environments. The library incorporates ergonomic features such as checking and waiting for server readiness for alignment, retrying different ports if the current one is occupied, and using an existing server if one has already been started.

We provide a variety of examples and documentation to help the user get started (Listing 1). Specifically, PxBLAT provides APIs, including Class *Server*, *Client*, and other free functions, to reproduce the BLAT suit. Class *Server* and *Client* have the same utilities as CLI *gfServer* and *gfClient*, respectively, but with greater flexibility. Free functions, for example, *start_server*, *query_server*, *status_server*, *fa2twobit*, *twobit2fa*, etc., serve as potential usage in different contexts. The library has been tested and developed with continuous integration (CI) and continuous development (CD) to ensure code quality. It uses type annotations for public classes and functions to guarantee quality and correctness via a type checker and static analyzer. The type annotations also make PxBLAT more user-friendly in developing environments, where the function signatures can be suggested and corrected automatically. Besides, PxBLAT contains CLIs implemented by its API. The CLI includes completion for different shells, improving versatility.

```python
from pxblat import Server
from pxblat import Client

client = Client(
    host="localhost",
    port=65000,
    seq_dir="ref/",
    min_score=20,
    min_identity=90,
)

server_option =
↪  Server.create_option().build()
with Server(
    host="localhost",
    port=65000,
    two_bit="ref/reference.2bit",
    option=server_option
) as server:
    work()  # work that consumes time
    server.wait_for_ready()
    result1 = client.query("ATCG")
    result2 = client.query("AtcG")
    result3 = client.query(["ATCG", "ATCG"])
    result4 = client.query(["cgTA",
    ↪  "fasta.fa"])

    for hsp in result1.hsps:
        print(f"{hsp}")
```

Listing 1. | **API Example.** The code snippet shows how to use the API of PxBLAT, and the query result can be iterated. More code examples can be found at `https://pxblat.readthedocs.io/en`

Benchmarking

The result and performance of PxBLAT are benchmarked against BLAT (V37.1) using a dataset of 1 FASTA Files. The testing dataset, which consists of five groups, is randomly sampled from chromosome 20 of the Human Genome (hg38), and each sample includes one sequence. The dataset includes varying lengths of sequence to cover a range of typical usage scenarios. The length of each sequence ranges from 1 bp to 3 bp (Figure S1). We summarized the high-scoring pairs (HSPs) of all hits in each sample derived from BLAT and PxBLAT and compared their predictions. We observed that PxBLAT reproduces the HSPs of BLAT across all tested samples, which ensures the correctness of PxBLAT (Table S1).

The benchmarking was conducted on an Apple M1 Pro with macOS 13.4.1 22F82 arm64. We use system calls to launch BLAT and the *time* library to measure the execution time. Generally, PxBLAT gains ~20 % speedup compared to BLAT (Table 1). In summary, PxBLAT offers tangible benefits according to reduced execution time and improved user experience, proving its value as an enhancement to BLAT's functionality.

Acknowledgements

Special thanks to the team who maintain the UCSC codebase and users from the bioinformatics community whose valuable feedback and suggestions were pivotal in refining PxBLAT's design and functionality.

Conflict of interest

None declared.

Funding

This work was supported by the National Institute of General Medical Sciences [R35GM142441].

bioRxiv preprint doi: https://doi.org/10.1101/2023.08.02.551686; this version posted October 24, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 1. Performance Benchmarking

Data Group	Samples (No.)	BLAT (Sec.)	PxBLAT (Sec.)	Speedup
benchmark/fas/1	200.00	61.40	50.41	1.22
benchmark/fas/2	200.00	61.70	51.10	1.21
benchmark/fas/3	200.00	55.34	46.44	1.19
benchmark/fas/4	200.00	56.98	47.72	1.19
benchmark/fas/5	200.00	54.55	46.44	1.17

Data availability

The dataset for benchmarking and testing is available at the GitHub repository `https://github.com/ylab-hi/pxblat`. The path of the dataset is *benchmark/fas*.

Code availability

The PxBLAT, along with the source code, is publicly available in the GitHub repository at `https://github.com/ylab-hi/pxblat`. The documentation is available at ReadtheDocs `https://pxblat.readthedocs.io/en/latest/`. The script for benchmarking is available at *tests/test_result.py* in the repository.

Reference

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. Basic local alignment search tool. 215(3):403–410. doi: 10.1016/S0022-2836(05)80360-2.

Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., and family=Hoon, p. u., given=Michiel J. L. Biopython: Freely available Python tools for computational molecular biology and bioinformatics. 25(11):1422–1423. doi: 10.1093/bioinformatics/btp163.

Higgins, D. G. and Sharp, P. M. CLUSTAL: A package for performing multiple sequence alignment on a microcomputer. 73(1):237–244. doi: 10.1016/0378-1119(88)90330-7.

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. 12(4):656–664. doi: 10.1101/gr.229202.

Perkel, J. M. Programming: Pick up Python. 518(7537):125–126. doi: 10.1038/518125a.

Putri, G. H., Anders, S., Pyl, P. T., Pimanda, J. E., and Zanini, F. Analysing high-throughput sequencing data in Python with HTSeq 2.0. 38(10):2943–2945. doi: 10.1093/bioinformatics/btac166.

bioRxiv preprint doi: https://doi.org/10.1101/2023.08.02.551686; this version posted October 24, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Information

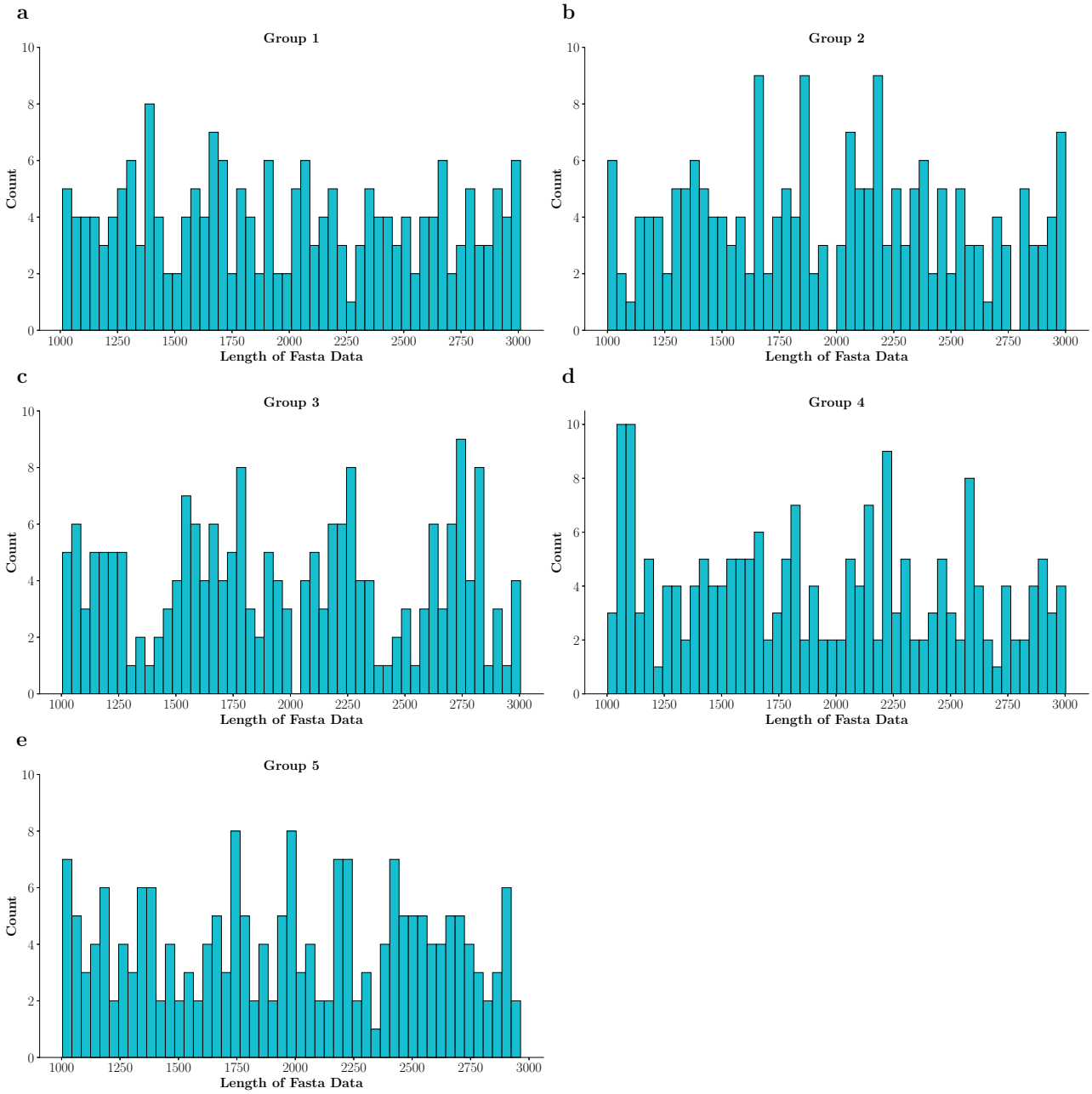

Figure S1. | **Distribution of Sequence Length.** The dataset for testing and benchmarking. Each group contains 200.00 fasta samples, and each fasta sample has one sequence. (**a**) The distribution of length in group 1 (**b**) The distribution of length in group 2 (**c**) The distribution of length in group 3 (**d**) The distribution of length in group 4 (**e**) The distribution of length in group 5

Table S1. Comparison of HSPs between BLAT and PxBLAT

Data	BLAT	PxBLAT
chr20-11648866-11650925	130.00	130.00
chr20-29850079-29852595	1.00	1.00
chr20-7493878-7496145	26.00	26.00
chr20-30152074-30153892	3.00	3.00
chr20-1140095-1141788	22.00	22.00

Continued on next page

bioRxiv preprint doi: https://doi.org/10.1101/2023.08.02.551686; this version posted October 24, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S1. Comparison of HSPs between BLAT and PxBLAT

Data	BLAT	PxBLAT
chr20-17520690-17523557	14.00	14.00
chr20-41374266-41376291	12.00	12.00
chr20-63240773-63243242	64.00	64.00
chr20-59286740-59288091	13.00	13.00
chr20-8061734-8063508	10.00	10.00
chr20-34429296-34431248	41.00	41.00
chr20-27977914-27979241	22.00	22.00
chr20-11695455-11697457	20.00	20.00
chr20-41792348-41794797	131.00	131.00
chr20-25858004-25860351	37.00	37.00
chr20-31404034-31407027	25.00	25.00
chr20-15852904-15855300	65.00	65.00
chr20-29917054-29919227	18.00	18.00
chr20-45175162-45177033	2.00	2.00
chr20-43937342-43939507	58.00	58.00
chr20-17073859-17076499	48.00	48.00
chr20-19823346-19824950	32.00	32.00
chr20-28295597-28298301	23.00	23.00
chr20-15826504-15829187	26.00	26.00
chr20-54683378-54684991	47.00	47.00
chr20-55475241-55476377	18.00	18.00
chr20-57949934-57952283	74.00	74.00
chr20-14442978-14444604	35.00	35.00
chr20-51293808-51295630	38.00	38.00
chr20-49194155-49196823	165.00	165.00
chr20-34696351-34697788	43.00	43.00
chr20-16044321-16045712	48.00	48.00
chr20-45309757-45312553	39.00	39.00
chr20-21063232-21065145	26.00	26.00
chr20-434138-436331	174.00	174.00
chr20-62110486-62112604	167.00	167.00
chr20-15866914-15868171	15.00	15.00
chr20-62343940-62345043	69.00	69.00
chr20-37168356-37170160	34.00	34.00
chr20-15699683-15701260	15.00	15.00
chr20-44826183-44827333	152.00	152.00
chr20-41538435-41540074	43.00	43.00
chr20-56032843-56034378	18.00	18.00
chr20-36813618-36816516	21.00	21.00
chr20-18637848-18640751	68.00	68.00
chr20-56294086-56296598	27.00	27.00
chr20-19514094-19515343	2.00	2.00
chr20-33893321-33895564	126.00	126.00
chr20-22282380-22283429	4.00	4.00
chr20-39910629-39912740	84.00	84.00
chr20-40898874-40901542	27.00	27.00
chr20-7536153-7537696	15.00	15.00
chr20-4818626-4819858	60.00	60.00
chr20-64316707-64319236	24.00	24.00
chr20-2960440-2961833	119.00	119.00
chr20-2519249-2521676	46.00	46.00
chr20-11596878-11599706	68.00	68.00
chr20-30829536-30831739	16.00	16.00
chr20-45188854-45190558	3.00	3.00

Continued on next page

bioRxiv preprint doi: https://doi.org/10.1101/2023.08.02.551686; this version posted October 24, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S1. Comparison of HSPs between BLAT and PxBLAT

Data	BLAT	PxBLAT
chr20-47238745-47241145	14.00	14.00
chr20-10538965-10540232	83.00	83.00
chr20-11996320-11998343	87.00	87.00
chr20-23440919-23442976	47.00	47.00
chr20-22451495-22453568	4.00	4.00
chr20-39709672-39711911	23.00	23.00
chr20-29782708-29784088	8.00	8.00
chr20-7911434-7913178	9.00	9.00
chr20-1508550-1511129	193.00	193.00
chr20-41451172-41452837	6.00	6.00
chr20-4003988-4006545	29.00	29.00
chr20-10921197-10923149	5.00	5.00
chr20-36789016-36790500	186.00	186.00
chr20-9309222-9310479	157.00	157.00
chr20-52532795-52534564	9.00	9.00
chr20-42081563-42082983	4.00	4.00
chr20-12605319-12606971	2.00	2.00
chr20-29389010-29390598	149.00	149.00
chr20-8128132-8130266	170.00	170.00
chr20-11399422-11401237	12.00	12.00
chr20-33251227-33252414	2.00	2.00
chr20-3342256-3343645	92.00	92.00
chr20-38869283-38872088	60.00	60.00
chr20-7649152-7650444	60.00	60.00
chr20-13021673-13023769	58.00	58.00
chr20-3971117-3972667	32.00	32.00
chr20-26232703-26234929	61.00	61.00
chr20-23063517-23064745	7.00	7.00
chr20-39079691-39080763	26.00	26.00
chr20-23882126-23884119	92.00	92.00
chr20-3170215-3172728	44.00	44.00
chr20-49421601-49423499	35.00	35.00
chr20-55380779-55383730	204.00	204.00
chr20-45794728-45796900	164.00	164.00
chr20-15159613-15161221	9.00	9.00
chr20-62378279-62380174	184.00	184.00
chr20-6934771-6936584	116.00	116.00
chr20-59538484-59541329	8.00	8.00
chr20-43661982-43664714	26.00	26.00
chr20-36707842-36709247	85.00	85.00
chr20-33898085-33900435	65.00	65.00
chr20-18062688-18064265	46.00	46.00
chr20-29490152-29491491	23.00	23.00
chr20-11837593-11840377	40.00	40.00
chr20-28934655-28937276	1.00	1.00
chr20-7689471-7691553	18.00	18.00
chr20-1746811-1749803	9.00	9.00
chr20-1261408-1262588	16.00	16.00
chr20-5757029-5759548	29.00	29.00
chr20-5361957-5364278	107.00	107.00
chr20-37305589-37308001	21.00	21.00
chr20-53702835-53705300	148.00	148.00
chr20-7447290-7449191	16.00	16.00
chr20-36626015-36629010	150.00	150.00

Continued on next page

bioRxiv preprint doi: https://doi.org/10.1101/2023.08.02.551686; this version posted October 24, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S1. Comparison of HSPs between BLAT and PxBLAT

Data	BLAT	PxBLAT
chr20-48184357-48185678	2.00	2.00
chr20-4634043-4636057	9.00	9.00
chr20-21292878-21295830	179.00	179.00
chr20-52551672-52553086	7.00	7.00
chr20-55410142-55411547	55.00	55.00
chr20-15164296-15166007	21.00	21.00
chr20-60294852-60297789	197.00	197.00
chr20-47112993-47114241	38.00	38.00
chr20-33645643-33648541	62.00	62.00
chr20-57099663-57101555	91.00	91.00
chr20-32443012-32444528	1.00	1.00
chr20-54597242-54598325	31.00	31.00
chr20-28168758-28171348	20.00	20.00
chr20-14255274-14256294	211.00	211.00
chr20-22858139-22859797	8.00	8.00
chr20-42977157-42978222	11.00	11.00
chr20-3890553-3891667	46.00	46.00
chr20-10043505-10044872	33.00	33.00
chr20-47292153-47293840	50.00	50.00
chr20-23406092-23408775	15.00	15.00
chr20-44245002-44246392	36.00	36.00
chr20-30999482-31002350	1.00	1.00
chr20-11990150-11991239	3.00	3.00
chr20-3810992-3812466	197.00	197.00
chr20-22091287-22094021	119.00	119.00
chr20-47477509-47479589	59.00	59.00
chr20-13750647-13753070	44.00	44.00
chr20-61483678-61484972	2.00	2.00
chr20-57584630-57586794	13.00	13.00
chr20-17904728-17906925	34.00	34.00
chr20-28086098-28087474	19.00	19.00
chr20-21317295-21318548	51.00	51.00
chr20-42758324-42759488	193.00	193.00
chr20-23580801-23581810	23.00	23.00
chr20-58929386-58930631	22.00	22.00
chr20-6605062-6607136	79.00	79.00
chr20-32987534-32989963	169.00	169.00
chr20-22493043-22494805	22.00	22.00
chr20-54190709-54192035	24.00	24.00
chr20-48343887-48346044	2.00	2.00
chr20-10262876-10265171	189.00	189.00
chr20-35374558-35376267	67.00	67.00
chr20-39018648-39021540	83.00	83.00
chr20-37785924-37788264	202.00	202.00
chr20-12575681-12578297	14.00	14.00
chr20-23374273-23375283	1.00	1.00
chr20-14278920-14280350	40.00	40.00
chr20-42778102-42780755	9.00	9.00
chr20-52810557-52811573	83.00	83.00
chr20-52564945-52566606	22.00	22.00
chr20-10598821-10601808	32.00	32.00
chr20-1172826-1175431	87.00	87.00
chr20-3347653-3350627	44.00	44.00
chr20-64176870-64179763	43.00	43.00

Continued on next page

bioRxiv preprint doi: https://doi.org/10.1101/2023.08.02.551686; this version posted October 24, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S1. Comparison of HSPs between BLAT and PxBLAT

Data	BLAT	PxBLAT
chr20-42829128-42830258	100.00	100.00
chr20-40623645-40625957	187.00	187.00
chr20-27009229-27011878	24.00	24.00
chr20-14732847-14734481	95.00	95.00
chr20-12698635-12699679	2.00	2.00
chr20-13328590-13330614	13.00	13.00
chr20-22912395-22915205	12.00	12.00
chr20-35339375-35342113	48.00	48.00
chr20-19431509-19433187	28.00	28.00
chr20-28727730-28729750	16.00	16.00
chr20-39515806-39517118	2.00	2.00
chr20-21857448-21859310	209.00	209.00
chr20-44751640-44753567	19.00	19.00
chr20-47371415-47374127	45.00	45.00
chr20-24140371-24142959	6.00	6.00
chr20-54083025-54084528	41.00	41.00
chr20-51775800-51777458	46.00	46.00
chr20-59400225-59402632	30.00	30.00
chr20-57144627-57147621	47.00	47.00
chr20-17690928-17692128	6.00	6.00
chr20-31238605-31240139	4.00	4.00
chr20-29259215-29262074	4.00	4.00
chr20-39132690-39134466	34.00	34.00
chr20-44674003-44676622	31.00	31.00
chr20-34435089-34436815	13.00	13.00
chr20-11085268-11087065	27.00	27.00
chr20-42010931-42012642	10.00	10.00
chr20-7483012-7484115	1.00	1.00
chr20-39873233-39875635	207.00	207.00

Table S2. Comparison of HSPs between BLAT and PxBLAT

Data	BLAT	PxBLAT
chr20-2828159-2830288	134.00	134.00
chr20-8971949-8973716	18.00	18.00
chr20-30095884-30097442	17.00	17.00
chr20-51719657-51722053	173.00	173.00
chr20-33974288-33975838	69.00	69.00
chr20-2278332-2280708	15.00	15.00
chr20-3395428-3397753	86.00	86.00
chr20-63803154-63805985	43.00	43.00
chr20-48031931-48034204	44.00	44.00
chr20-46308727-46309965	6.00	6.00
chr20-14083167-14084282	167.00	167.00
chr20-45353941-45356038	53.00	53.00
chr20-63851619-63854055	64.00	64.00
chr20-50583738-50585129	3.00	3.00
chr20-13751249-13752454	42.00	42.00
chr20-49452302-49454684	159.00	159.00
chr20-18383256-18384901	73.00	73.00
chr20-29651400-29652794	9.00	9.00
chr20-40330327-40331786	18.00	18.00
chr20-35818754-35821105	100.00	100.00

Continued on next page

bioRxiv preprint doi: https://doi.org/10.1101/2023.08.02.551686; this version posted October 24, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S2. Comparison of HSPs between BLAT and PxBLAT

Data	BLAT	PxBLAT
chr20-2688118-2690258	39.00	39.00
chr20-47399940-47402884	114.00	114.00
chr20-50899780-50901977	19.00	19.00
chr20-43133225-43135909	192.00	192.00
chr20-46544773-46547501	10.00	10.00
chr20-19754747-19756349	18.00	18.00
chr20-223022-225549	9.00	9.00
chr20-63046239-63049189	47.00	47.00
chr20-57304541-57306258	58.00	58.00
chr20-38493743-38495091	25.00	25.00
chr20-20265635-20267968	16.00	16.00
chr20-2844781-2847151	93.00	93.00
chr20-53725027-53727516	36.00	36.00
chr20-10920657-10922515	8.00	8.00
chr20-44371315-44374295	39.00	39.00
chr20-3431376-3433251	202.00	202.00
chr20-39135488-39137700	52.00	52.00
chr20-58745286-58747067	11.00	11.00
chr20-31791374-31793816	205.00	205.00
chr20-36971580-36972630	27.00	27.00
chr20-45741446-45744259	26.00	26.00
chr20-28837189-28840052	1.00	1.00
chr20-53642354-53643485	103.00	103.00
chr20-63846146-63848461	16.00	16.00
chr20-24600426-24602535	54.00	54.00
chr20-12493566-12496209	178.00	178.00
chr20-24425863-24428862	24.00	24.00
chr20-43710550-43713461	126.00	126.00
chr20-3694835-3697763	73.00	73.00
chr20-49087699-49089567	8.00	8.00
chr20-51447199-51448636	88.00	88.00
chr20-8640913-8642978	6.00	6.00
chr20-53874690-53877592	56.00	56.00
chr20-32558083-32560533	37.00	37.00
chr20-39866491-39868448	32.00	32.00
chr20-57235806-57237280	2.00	2.00
chr20-1551870-1553269	2.00	2.00
chr20-60276391-60278170	31.00	31.00
chr20-47735701-47737760	12.00	12.00
chr20-47950200-47952033	26.00	26.00
chr20-37264292-37266206	95.00	95.00
chr20-25972909-25975750	5.00	5.00
chr20-49810048-49811680	22.00	22.00
chr20-7730740-7733285	45.00	45.00
chr20-51212931-51215436	46.00	46.00
chr20-35679220-35680810	178.00	178.00
chr20-60278816-60280383	4.00	4.00
chr20-1848711-1850476	89.00	89.00
chr20-61487368-61489909	64.00	64.00
chr20-37154283-37155740	35.00	35.00
chr20-10295968-10296987	7.00	7.00
chr20-61931058-61933207	7.00	7.00
chr20-59480418-59482175	191.00	191.00
chr20-47105521-47106711	109.00	109.00

Continued on next page

bioRxiv preprint doi: https://doi.org/10.1101/2023.08.02.551686; this version posted October 24, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S2. Comparison of HSPs between BLAT and PxBLAT

Data	BLAT	PxBLAT
chr20-28467685-28469748	20.00	20.00
chr20-23363427-23366158	26.00	26.00
chr20-1681181-1682213	28.00	28.00
chr20-23846256-23847406	31.00	31.00
chr20-55274707-55276347	223.00	223.00
chr20-63347021-63349649	4.00	4.00
chr20-40080433-40082317	75.00	75.00
chr20-9530092-9532401	1.00	1.00
chr20-38108034-38110387	56.00	56.00
chr20-9623861-9625540	5.00	5.00
chr20-58197693-58200651	48.00	48.00
chr20-14255898-14258880	13.00	13.00
chr20-29162172-29164046	18.00	18.00
chr20-3651661-3653661	19.00	19.00
chr20-54221532-54223373	24.00	24.00
chr20-30550507-30552616	3.00	3.00
chr20-54778098-54779435	3.00	3.00
chr20-34554504-34557124	195.00	195.00
chr20-46153137-46156011	61.00	61.00
chr20-19287160-19289355	61.00	61.00
chr20-15370716-15372514	6.00	6.00
chr20-12817209-12820043	213.00	213.00
chr20-26494131-26495614	26.00	26.00
chr20-44385075-44387095	34.00	34.00
chr20-30229549-30231195	216.00	216.00
chr20-58407287-58410250	54.00	54.00
chr20-22419707-22421972	5.00	5.00
chr20-64189992-64192899	44.00	44.00
chr20-60602761-60605519	57.00	57.00
chr20-29075089-29077901	22.00	22.00
chr20-51434881-51436198	39.00	39.00
chr20-29854637-29855989	3.00	3.00
chr20-29127537-29128926	16.00	16.00
chr20-3190191-3192877	35.00	35.00
chr20-35916578-35918313	59.00	59.00
chr20-56918427-56920492	34.00	34.00
chr20-41848804-41850751	2.00	2.00
chr20-26793231-26795382	22.00	22.00
chr20-64014289-64015618	42.00	42.00
chr20-39382394-39384063	6.00	6.00
chr20-11297686-11298759	29.00	29.00
chr20-53600600-53601859	25.00	25.00
chr20-7664191-7666587	117.00	117.00
chr20-50657586-50658971	43.00	43.00
chr20-14086263-14087689	27.00	27.00
chr20-31644053-31645876	202.00	202.00
chr20-49881247-49883414	19.00	19.00
chr20-39045849-39048318	71.00	71.00
chr20-26667702-26668733	24.00	24.00
chr20-35794840-35797816	33.00	33.00
chr20-14041285-14043655	43.00	43.00
chr20-43887719-43889010	77.00	77.00
chr20-48557745-48559863	75.00	75.00
chr20-13636473-13637790	22.00	22.00

Continued on next page

bioRxiv preprint doi: https://doi.org/10.1101/2023.08.02.551686; this version posted October 24, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S2. Comparison of HSPs between BLAT and PxBLAT

Data	BLAT	PxBLAT
chr20-60386487-60388629	85.00	85.00
chr20-46602679-46604724	105.00	105.00
chr20-6342124-6344048	2.00	2.00
chr20-52509804-52511365	24.00	24.00
chr20-8596729-8597957	184.00	184.00
chr20-46229584-46231629	10.00	10.00
chr20-32426330-32428027	49.00	49.00
chr20-59309503-59310835	26.00	26.00
chr20-44725323-44726823	169.00	169.00
chr20-3932045-3934068	4.00	4.00
chr20-923271-925546	23.00	23.00
chr20-8844646-8845932	118.00	118.00
chr20-16341258-16343778	11.00	11.00
chr20-11253087-11254915	6.00	6.00
chr20-62299195-62300925	16.00	16.00
chr20-13574193-13576883	36.00	36.00
chr20-47318361-47320801	67.00	67.00
chr20-38097496-38099695	110.00	110.00
chr20-16852028-16854617	9.00	9.00
chr20-48876739-48878931	38.00	38.00
chr20-46598465-46599826	9.00	9.00
chr20-50733724-50735913	31.00	31.00
chr20-16694673-16696959	41.00	41.00
chr20-60554396-60556059	3.00	3.00
chr20-7100666-7102477	22.00	22.00
chr20-9052261-9055087	202.00	202.00
chr20-47716012-47718585	49.00	49.00
chr20-55091558-55093023	18.00	18.00
chr20-55647967-55649485	37.00	37.00
chr20-12070860-12072133	201.00	201.00
chr20-46040466-46042741	17.00	17.00
chr20-59258885-59260725	58.00	58.00
chr20-41615264-41617426	6.00	6.00
chr20-47812008-47813426	32.00	32.00
chr20-53173457-53174462	31.00	31.00
chr20-47785182-47786828	10.00	10.00
chr20-17404964-17407937	2.00	2.00
chr20-33100162-33102029	19.00	19.00
chr20-38669941-38671179	89.00	89.00
chr20-24203548-24204576	1.00	1.00
chr20-56658618-56660794	26.00	26.00
chr20-9502811-9505031	39.00	39.00
chr20-444175-446235	55.00	55.00
chr20-64042363-64044772	66.00	66.00
chr20-42146908-42148554	5.00	5.00
chr20-31814615-31816841	14.00	14.00
chr20-7185589-7186727	23.00	23.00
chr20-4210142-4212478	108.00	108.00
chr20-2738418-2740895	40.00	40.00
chr20-24397070-24398500	172.00	172.00
chr20-62308298-62309851	1.00	1.00
chr20-25326590-25328009	35.00	35.00
chr20-16727307-16729492	45.00	45.00
chr20-38178690-38181407	26.00	26.00

Continued on next page

bioRxiv preprint doi: https://doi.org/10.1101/2023.08.02.551686; this version posted October 24, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S2. Comparison of HSPs between BLAT and PxBLAT

Data	BLAT	PxBLAT
chr20-23745409-23746596	4.00	4.00
chr20-28209126-28211757	22.00	22.00
chr20-14538504-14540773	3.00	3.00
chr20-18389195-18390366	1.00	1.00
chr20-48698281-48699926	90.00	90.00
chr20-17745800-17748778	21.00	21.00
chr20-17202460-17203460	18.00	18.00
chr20-3858012-3859518	62.00	62.00
chr20-38806155-38807719	5.00	5.00
chr20-59941646-59943512	10.00	10.00
chr20-18936575-18939153	12.00	12.00
chr20-45510964-45512824	41.00	41.00
chr20-26624053-26625245	21.00	21.00
chr20-37191072-37192377	38.00	38.00

Table S3. Comparison of HSPs between BLAT and PxBLAT

Data	BLAT	PxBLAT
chr20-32029134-32032031	93.00	93.00
chr20-38052616-38054138	32.00	32.00
chr20-2602976-2605324	19.00	19.00
chr20-42169915-42172757	15.00	15.00
chr20-62125729-62128471	19.00	19.00
chr20-2162753-2164165	5.00	5.00
chr20-34630972-34632436	113.00	113.00
chr20-6611620-6614440	152.00	152.00
chr20-19300360-19302642	6.00	6.00
chr20-25189802-25191365	154.00	154.00
chr20-15951964-15953450	26.00	26.00
chr20-47285001-47287102	52.00	52.00
chr20-22548369-22549974	9.00	9.00
chr20-24725480-24727693	52.00	52.00
chr20-55633941-55636718	5.00	5.00
chr20-30324027-30326285	30.00	30.00
chr20-34789703-34791591	27.00	27.00
chr20-22886111-22887126	5.00	5.00
chr20-52373722-52374925	52.00	52.00
chr20-62594615-62596270	61.00	61.00
chr20-23567798-23569994	5.00	5.00
chr20-52746782-52748970	26.00	26.00
chr20-63919107-63920159	20.00	20.00
chr20-37603411-37606081	21.00	21.00
chr20-45549091-45551404	167.00	167.00
chr20-24122821-24124880	2.00	2.00
chr20-39384970-39386746	47.00	47.00
chr20-56953750-56954962	11.00	11.00
chr20-27265832-27268665	27.00	27.00
chr20-989008-991475	96.00	96.00
chr20-13233506-13234845	9.00	9.00
chr20-54079315-54081921	56.00	56.00
chr20-22276591-22279281	202.00	202.00
chr20-32344273-32347087	26.00	26.00
chr20-14761627-14763228	15.00	15.00

Continued on next page

bioRxiv preprint doi: https://doi.org/10.1101/2023.08.02.551686; this version posted October 24, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S3. Comparison of HSPs between BLAT and PxBLAT

Data	BLAT	PxBLAT
chr20-24471803-24474607	7.00	7.00
chr20-22305788-22308754	43.00	43.00
chr20-63927315-63929626	54.00	54.00
chr20-33534721-33537439	158.00	158.00
chr20-53020990-53022920	33.00	33.00
chr20-44701412-44703325	34.00	34.00
chr20-63523339-63524524	190.00	190.00
chr20-39750571-39752199	50.00	50.00
chr20-47879542-47880820	162.00	162.00
chr20-12949055-12950079	1.00	1.00
chr20-37737796-37740734	26.00	26.00
chr20-17600492-17603407	68.00	68.00
chr20-40848143-40849212	2.00	2.00
chr20-33135173-33136880	131.00	131.00
chr20-39118192-39119260	4.00	4.00
chr20-27690465-27692366	23.00	23.00
chr20-37858061-37860321	49.00	49.00
chr20-12405605-12407882	6.00	6.00
chr20-57386121-57387852	70.00	70.00
chr20-34902889-34905611	39.00	39.00
chr20-52070637-52073331	35.00	35.00
chr20-40779660-40781751	47.00	47.00
chr20-31098450-31100717	3.00	3.00
chr20-29969458-29971235	18.00	18.00
chr20-13789470-13791842	152.00	152.00
chr20-26814485-26817257	22.00	22.00
chr20-40740691-40742983	29.00	29.00
chr20-49955693-49958425	183.00	183.00
chr20-37813305-37814584	9.00	9.00
chr20-18255943-18258048	64.00	64.00
chr20-63109432-63111043	34.00	34.00
chr20-14211039-14213272	9.00	9.00
chr20-17100441-17103140	9.00	9.00
chr20-2866996-2869328	38.00	38.00
chr20-53703835-53705398	158.00	158.00
chr20-51431118-51433743	56.00	56.00
chr20-6192786-6193790	196.00	196.00
chr20-10748356-10751019	35.00	35.00
chr20-33974548-33976407	44.00	44.00
chr20-32270385-32273111	20.00	20.00
chr20-36902282-36904506	28.00	28.00
chr20-198832-201190	40.00	40.00
chr20-7927314-7928574	27.00	27.00
chr20-17899461-17900502	58.00	58.00
chr20-11120841-11122366	23.00	23.00
chr20-59407661-59409193	27.00	27.00
chr20-35649824-35652625	5.00	5.00
chr20-30688628-30690580	48.00	48.00
chr20-49466264-49467309	5.00	5.00
chr20-63692672-63694078	2.00	2.00
chr20-12570747-12573421	5.00	5.00
chr20-8512304-8514491	20.00	20.00
chr20-24204303-24205314	2.00	2.00
chr20-11639430-11641175	20.00	20.00

Continued on next page

bioRxiv preprint doi: https://doi.org/10.1101/2023.08.02.551686; this version posted October 24, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S3. Comparison of HSPs between BLAT and PxBLAT

Data	BLAT	PxBLAT
chr20-26938330-26940788	24.00	24.00
chr20-17587799-17589501	209.00	209.00
chr20-26384682-26386149	23.00	23.00
chr20-46308421-46310414	41.00	41.00
chr20-44079162-44081096	23.00	23.00
chr20-21475551-21476686	50.00	50.00
chr20-13972488-13974627	44.00	44.00
chr20-45697903-45699479	88.00	88.00
chr20-24838480-24839925	11.00	11.00
chr20-7742386-7745111	4.00	4.00
chr20-56522851-56524815	12.00	12.00
chr20-22497057-22499773	16.00	16.00
chr20-51998729-52000293	22.00	22.00
chr20-38767581-38768689	5.00	5.00
chr20-19863311-19865822	56.00	56.00
chr20-9295230-9297764	226.00	226.00
chr20-15012779-15015584	8.00	8.00
chr20-5041846-5044587	56.00	56.00
chr20-580210-582022	35.00	35.00
chr20-15328579-15331324	42.00	42.00
chr20-49481967-49483214	17.00	17.00
chr20-5260817-5262044	5.00	5.00
chr20-23340929-23343084	102.00	102.00
chr20-2269716-2270774	189.00	189.00
chr20-46399947-46401667	31.00	31.00
chr20-10815100-10817410	32.00	32.00
chr20-45864818-45866440	53.00	53.00
chr20-45489949-45492069	35.00	35.00
chr20-21302718-21304312	1.00	1.00
chr20-9931834-9933992	5.00	5.00
chr20-56295194-56296721	5.00	5.00
chr20-38225781-38227971	18.00	18.00
chr20-18585659-18587381	35.00	35.00
chr20-16940912-16942721	39.00	39.00
chr20-3758200-3759977	4.00	4.00
chr20-11807056-11808124	3.00	3.00
chr20-127575-129488	93.00	93.00
chr20-40755444-40756701	5.00	5.00
chr20-21011682-21013735	2.00	2.00
chr20-10119043-10120171	59.00	59.00
chr20-1313128-1315640	18.00	18.00
chr20-11712977-11715178	142.00	142.00
chr20-34597997-34600726	28.00	28.00
chr20-53048517-53050101	29.00	29.00
chr20-15161566-15163488	8.00	8.00
chr20-15137364-15139959	203.00	203.00
chr20-60342420-60345056	94.00	94.00
chr20-55169861-55170982	156.00	156.00
chr20-24314101-24316265	27.00	27.00
chr20-53183314-53186148	16.00	16.00
chr20-43887385-43889661	75.00	75.00
chr20-59470391-59471545	1.00	1.00
chr20-53159093-53161302	35.00	35.00
chr20-5895129-5896312	26.00	26.00

Continued on next page

bioRxiv preprint doi: https://doi.org/10.1101/2023.08.02.551686; this version posted October 24, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S3. Comparison of HSPs between BLAT and PxBLAT

Data	BLAT	PxBLAT
chr20-25292499-25293803	3.00	3.00
chr20-34161145-34162369	41.00	41.00
chr20-45868564-45870103	74.00	74.00
chr20-4887453-4889888	3.00	3.00
chr20-39559840-39561722	9.00	9.00
chr20-33221436-33223116	56.00	56.00
chr20-755464-758088	35.00	35.00
chr20-1111176-1112831	46.00	46.00
chr20-2802607-2803838	36.00	36.00
chr20-38997152-38999498	22.00	22.00
chr20-5049946-5052675	35.00	35.00
chr20-48767423-48769196	48.00	48.00
chr20-2659893-2661452	56.00	56.00
chr20-49121260-49124086	21.00	21.00
chr20-54247731-54250009	75.00	75.00
chr20-8334685-8337480	34.00	34.00
chr20-1762478-1764560	46.00	46.00
chr20-53286968-53289954	23.00	23.00
chr20-23613637-23615158	53.00	53.00
chr20-7135001-7136753	88.00	88.00
chr20-59872605-59875175	5.00	5.00
chr20-33944307-33946354	27.00	27.00
chr20-29344297-29347205	8.00	8.00
chr20-28334593-28336360	22.00	22.00
chr20-49230825-49232042	66.00	66.00
chr20-48044237-48045631	3.00	3.00
chr20-26227485-26229228	5.00	5.00
chr20-26487906-26489572	25.00	25.00
chr20-11596974-11599728	65.00	65.00
chr20-45661385-45663135	194.00	194.00
chr20-23562380-23564047	2.00	2.00
chr20-5735736-5738602	12.00	12.00
chr20-55535504-55537712	3.00	3.00
chr20-45493939-45496533	99.00	99.00
chr20-25430083-25432571	107.00	107.00
chr20-2981029-2982513	77.00	77.00
chr20-49523722-49525938	32.00	32.00
chr20-27912294-27914399	20.00	20.00
chr20-15366252-15368022	33.00	33.00
chr20-40846798-40849784	29.00	29.00
chr20-20650550-20653192	59.00	59.00
chr20-13545370-13547154	2.00	2.00
chr20-27596727-27598551	19.00	19.00
chr20-3394620-3396590	171.00	171.00
chr20-1444330-1446279	3.00	3.00
chr20-42147101-42148224	2.00	2.00
chr20-60359901-60361245	11.00	11.00
chr20-58137613-58139410	3.00	3.00
chr20-40224761-40225920	10.00	10.00
chr20-57324246-57325405	13.00	13.00
chr20-21713213-21715491	39.00	39.00
chr20-14037568-14040191	83.00	83.00
chr20-16385063-16388039	118.00	118.00
chr20-59768230-59769419	5.00	5.00

bioRxiv preprint doi: https://doi.org/10.1101/2023.08.02.551686; this version posted October 24, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S4. Comparison of HSPs between BLAT and PxBLAT

Data	BLAT	PxBLAT
chr20-31727237-31729024	39.00	39.00
chr20-36152090-36154966	151.00	151.00
chr20-1018906-1020553	11.00	11.00
chr20-9821106-9823172	75.00	75.00
chr20-52034533-52037144	85.00	85.00
chr20-15008181-15010280	66.00	66.00
chr20-41851737-41853531	6.00	6.00
chr20-29408974-29411107	20.00	20.00
chr20-10955881-10957008	1.00	1.00
chr20-8970270-8971572	18.00	18.00
chr20-2350436-2352589	200.00	200.00
chr20-35494115-35496325	37.00	37.00
chr20-16994461-16996880	203.00	203.00
chr20-1121638-1123536	20.00	20.00
chr20-10539946-10542788	76.00	76.00
chr20-31981394-31984035	24.00	24.00
chr20-31059406-31062015	9.00	9.00
chr20-23005241-23006490	85.00	85.00
chr20-27678668-27680740	22.00	22.00
chr20-23089029-23090440	1.00	1.00
chr20-18766639-18768846	74.00	74.00
chr20-35270491-35271927	24.00	24.00
chr20-32671767-32673927	46.00	46.00
chr20-5469364-5470405	1.00	1.00
chr20-44586665-44588074	3.00	3.00
chr20-19751371-19752469	50.00	50.00
chr20-11475039-11477338	4.00	4.00
chr20-52930875-52933353	52.00	52.00
chr20-23941121-23942933	189.00	189.00
chr20-10612667-10615654	16.00	16.00
chr20-39220431-39222072	6.00	6.00
chr20-24056618-24059198	190.00	190.00
chr20-63753892-63756615	196.00	196.00
chr20-64298054-64301011	16.00	16.00
chr20-43632072-43634462	46.00	46.00
chr20-29402266-29403661	49.00	49.00
chr20-34611441-34612505	39.00	39.00
chr20-16219627-16222599	19.00	19.00
chr20-47623287-47624373	18.00	18.00
chr20-37572653-37574987	20.00	20.00
chr20-57328038-57329145	159.00	159.00
chr20-54255861-54257299	11.00	11.00
chr20-17523656-17525455	3.00	3.00
chr20-3054439-3055812	26.00	26.00
chr20-51095602-51096740	8.00	8.00
chr20-26295259-26296843	6.00	6.00
chr20-16956946-16959866	4.00	4.00
chr20-59871467-59873730	18.00	18.00
chr20-12972124-12973571	9.00	9.00
chr20-64187505-64190371	7.00	7.00
chr20-20036801-20038441	54.00	54.00
chr20-52361579-52363650	49.00	49.00
chr20-48763637-48765211	21.00	21.00
chr20-29272745-29274078	3.00	3.00

Continued on next page

bioRxiv preprint doi: https://doi.org/10.1101/2023.08.02.551686; this version posted October 24, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S4. Comparison of HSPs between BLAT and PxBLAT

Data	BLAT	PxBLAT
chr20-6596603-6597647	1.00	1.00
chr20-25092945-25094053	38.00	38.00
chr20-23976043-23978934	34.00	34.00
chr20-33528290-33530524	183.00	183.00
chr20-45968564-45971166	41.00	41.00
chr20-29195839-29197450	17.00	17.00
chr20-12991634-12993772	18.00	18.00
chr20-7287285-7289523	45.00	45.00
chr20-12780972-12782351	20.00	20.00
chr20-17273569-17275885	46.00	46.00
chr20-53404329-53405379	8.00	8.00
chr20-32199416-32201827	159.00	159.00
chr20-56899277-56900938	64.00	64.00
chr20-21026086-21027263	198.00	198.00
chr20-29495766-29497356	5.00	5.00
chr20-46509725-46510768	176.00	176.00
chr20-44370323-44372070	36.00	36.00
chr20-34807421-34810326	128.00	128.00
chr20-29310629-29312542	80.00	80.00
chr20-1861413-1862970	10.00	10.00
chr20-13425759-13427299	2.00	2.00
chr20-40797946-40799834	14.00	14.00
chr20-9134988-9136715	38.00	38.00
chr20-6129194-6130506	20.00	20.00
chr20-32328850-32331341	40.00	40.00
chr20-37816931-37818838	30.00	30.00
chr20-43643216-43645801	82.00	82.00
chr20-57951819-57954389	39.00	39.00
chr20-38495333-38497995	69.00	69.00
chr20-23579562-23580734	13.00	13.00
chr20-32476687-32477738	89.00	89.00
chr20-46657968-46660258	17.00	17.00
chr20-44029371-44030373	1.00	1.00
chr20-19040746-19042992	24.00	24.00
chr20-41547378-41548428	5.00	5.00
chr20-31338848-31340507	33.00	33.00
chr20-50733531-50736468	30.00	30.00
chr20-23264110-23266210	46.00	46.00
chr20-59975068-59976189	25.00	25.00
chr20-40116512-40117628	2.00	2.00
chr20-34348380-34351135	191.00	191.00
chr20-53857114-53859581	90.00	90.00
chr20-24510328-24512009	40.00	40.00
chr20-2400648-2401902	51.00	51.00
chr20-41951417-41953958	8.00	8.00
chr20-10519505-10521311	173.00	173.00
chr20-38229341-38232239	39.00	39.00
chr20-1961461-1962512	4.00	4.00
chr20-4493702-4496293	164.00	164.00
chr20-39058190-39059640	65.00	65.00
chr20-11840010-11842763	159.00	159.00
chr20-51385931-51387697	2.00	2.00
chr20-32743241-32746228	61.00	61.00
chr20-36043748-36044868	4.00	4.00

Continued on next page

bioRxiv preprint doi: https://doi.org/10.1101/2023.08.02.551686; this version posted October 24, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S4. Comparison of HSPs between BLAT and PxBLAT

Data	BLAT	PxBLAT
chr20-47940674-47941937	27.00	27.00
chr20-22906294-22908788	8.00	8.00
chr20-46563742-46565560	42.00	42.00
chr20-34970963-34972526	22.00	22.00
chr20-46825075-46827145	11.00	11.00
chr20-6428315-6430827	188.00	188.00
chr20-42217777-42219934	16.00	16.00
chr20-19913134-19914585	34.00	34.00
chr20-37718004-37720318	152.00	152.00
chr20-51873445-51875614	85.00	85.00
chr20-21185966-21187683	41.00	41.00
chr20-36715891-36717659	71.00	71.00
chr20-25706788-25709614	123.00	123.00
chr20-50918346-50919658	55.00	55.00
chr20-19216584-19217715	110.00	110.00
chr20-60913180-60914343	27.00	27.00
chr20-40445916-40447302	2.00	2.00
chr20-39471792-39473038	200.00	200.00
chr20-36763453-36764650	82.00	82.00
chr20-58906803-58909272	170.00	170.00
chr20-24358042-24359864	9.00	9.00
chr20-5054910-5055935	113.00	113.00
chr20-39718409-39719767	60.00	60.00
chr20-48995874-48997427	47.00	47.00
chr20-11060084-11062288	61.00	61.00
chr20-3041121-3043902	46.00	46.00
chr20-11729530-11731335	4.00	4.00
chr20-6876021-6878545	44.00	44.00
chr20-60459034-60461357	22.00	22.00
chr20-50558110-50560711	68.00	68.00
chr20-1551191-1553332	2.00	2.00
chr20-23189149-23191563	50.00	50.00
chr20-46682255-46683957	5.00	5.00
chr20-46693319-46694364	56.00	56.00
chr20-50170226-50172189	11.00	11.00
chr20-63698924-63701762	5.00	5.00
chr20-62756096-62757580	69.00	69.00
chr20-55371351-55373202	30.00	30.00
chr20-61650045-61652414	3.00	3.00
chr20-33091750-33093858	24.00	24.00
chr20-61216684-61218692	7.00	7.00
chr20-47978091-47980293	85.00	85.00
chr20-16870384-16872626	14.00	14.00
chr20-55265469-55267149	6.00	6.00
chr20-42713744-42715938	38.00	38.00
chr20-20608914-20611491	36.00	36.00
chr20-59966328-59967826	57.00	57.00
chr20-29829808-29831310	8.00	8.00
chr20-55321262-55323731	34.00	34.00
chr20-38905281-38906801	57.00	57.00
chr20-5081835-5084428	35.00	35.00
chr20-12399681-12402397	19.00	19.00
chr20-15787670-15789236	10.00	10.00
chr20-61849267-61851471	112.00	112.00

Continued on next page

bioRxiv preprint doi: https://doi.org/10.1101/2023.08.02.551686; this version posted October 24, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S4. Comparison of HSPs between BLAT and PxBLAT

Data	BLAT	PxBLAT
chr20-42224821-42226456	8.00	8.00
chr20-2200973-2202184	2.00	2.00
chr20-44016368-44018208	138.00	138.00
chr20-42732744-42735046	2.00	2.00
chr20-8901543-8903090	37.00	37.00
chr20-33361770-33363718	16.00	16.00
chr20-60319529-60322435	17.00	17.00
chr20-9417083-9419320	19.00	19.00
chr20-47368389-47371245	181.00	181.00
chr20-49829087-49831301	97.00	97.00
chr20-45577205-45579659	30.00	30.00
chr20-9885847-9887724	4.00	4.00
chr20-34086336-34087450	25.00	25.00
chr20-12690873-12693547	25.00	25.00
chr20-61265483-61267506	31.00	31.00
chr20-60767473-60770234	21.00	21.00
chr20-37096290-37097735	46.00	46.00
chr20-19806699-19808018	39.00	39.00
chr20-5383554-5385658	41.00	41.00
chr20-46475786-46476831	62.00	62.00
chr20-35346166-35347602	30.00	30.00
chr20-39431060-39432816	28.00	28.00
chr20-50643697-50644883	205.00	205.00
chr20-22403148-22404794	15.00	15.00
chr20-26331430-26334030	3.00	3.00
chr20-14375436-14376511	196.00	196.00
chr20-32839818-32840903	124.00	124.00
chr20-20727967-20730763	59.00	59.00
chr20-25447522-25449068	21.00	21.00
chr20-52094840-52096680	143.00	143.00
chr20-21563954-21565882	146.00	146.00
chr20-5436550-5437659	70.00	70.00
chr20-63482588-63484649	191.00	191.00
chr20-19772825-19774949	43.00	43.00
chr20-38393781-38395769	19.00	19.00
chr20-34751103-34754059	175.00	175.00
chr20-26987677-26990651	24.00	24.00
chr20-8530344-8531955	41.00	41.00

Table S5. Comparison of HSPs between BLAT and PxBLAT

Data	BLAT	PxBLAT
chr20-7777117-7779958	31.00	31.00
chr20-28090646-28093037	21.00	21.00
chr20-61013940-61015279	85.00	85.00
chr20-3960118-3961869	55.00	55.00
chr20-39365878-39368085	195.00	195.00
chr20-18796067-18798923	20.00	20.00
chr20-46223156-46225144	57.00	57.00
chr20-61459668-61462165	18.00	18.00
chr20-1880831-1883521	76.00	76.00
chr20-61691915-61692999	85.00	85.00
chr20-16641272-16643484	17.00	17.00

Continued on next page

bioRxiv preprint doi: https://doi.org/10.1101/2023.08.02.551686; this version posted October 24, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S5. Comparison of HSPs between BLAT and PxBLAT

Data	BLAT	PxBLAT
chr20-53437738-53438813	25.00	25.00
chr20-2921736-2924355	36.00	36.00
chr20-9786768-9787896	83.00	83.00
chr20-3348853-3349856	52.00	52.00
chr20-26665706-26668235	23.00	23.00
chr20-36357607-36360008	122.00	122.00
chr20-28708909-28711108	16.00	16.00
chr20-55053702-55056597	17.00	17.00
chr20-12411633-12414319	6.00	6.00
chr20-58261678-58263991	191.00	191.00
chr20-31909700-31910874	18.00	18.00
chr20-13559545-13561467	38.00	38.00
chr20-41522731-41524518	71.00	71.00
chr20-16832380-16833904	27.00	27.00
chr20-273245-275299	34.00	34.00
chr20-45248820-45251295	17.00	17.00
chr20-12906343-12908335	33.00	33.00
chr20-63453750-63455013	6.00	6.00
chr20-2471058-2473777	50.00	50.00
chr20-37782559-37785511	10.00	10.00
chr20-32511957-32514426	19.00	19.00
chr20-2019303-2021924	197.00	197.00
chr20-18049373-18051341	47.00	47.00
chr20-12831320-12833476	15.00	15.00
chr20-21527118-21529684	34.00	34.00
chr20-60252763-60255169	195.00	195.00
chr20-30750577-30753179	25.00	25.00
chr20-60065233-60067754	22.00	22.00
chr20-22659147-22660560	2.00	2.00
chr20-59754388-59756018	4.00	4.00
chr20-53368185-53369917	25.00	25.00
chr20-58016931-58019589	14.00	14.00
chr20-36346997-36349477	78.00	78.00
chr20-3567642-3569619	110.00	110.00
chr20-19249962-19251073	7.00	7.00
chr20-7460107-7461273	119.00	119.00
chr20-41996994-41998017	2.00	2.00
chr20-38953285-38954932	39.00	39.00
chr20-15693304-15695054	12.00	12.00
chr20-49975829-49978524	30.00	30.00
chr20-48174502-48175954	2.00	2.00
chr20-57547413-57549744	12.00	12.00
chr20-27832062-27834490	23.00	23.00
chr20-13033294-13035107	10.00	10.00
chr20-25043139-25044218	7.00	7.00
chr20-58699090-58701273	22.00	22.00
chr20-33407023-33408027	161.00	161.00
chr20-48360301-48362055	199.00	199.00
chr20-48967663-48970119	174.00	174.00
chr20-53747260-53749122	37.00	37.00
chr20-55195931-55198585	12.00	12.00
chr20-47766917-47769131	51.00	51.00
chr20-35692775-35694627	57.00	57.00
chr20-61078167-61079319	2.00	2.00

Continued on next page

bioRxiv preprint doi: https://doi.org/10.1101/2023.08.02.551686; this version posted October 24, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S5. Comparison of HSPs between BLAT and PxBLAT

Data	BLAT	PxBLAT
chr20-45472353-45474572	210.00	210.00
chr20-5253614-5255621	20.00	20.00
chr20-27113258-27114615	24.00	24.00
chr20-25345213-25347164	112.00	112.00
chr20-20789302-20792218	112.00	112.00
chr20-12282317-12283946	17.00	17.00
chr20-16112151-16113224	16.00	16.00
chr20-63334929-63336520	33.00	33.00
chr20-38100003-38101534	37.00	37.00
chr20-1557202-1559796	38.00	38.00
chr20-43180645-43183274	21.00	21.00
chr20-61948125-61949374	2.00	2.00
chr20-15302550-15304778	4.00	4.00
chr20-468292-469632	4.00	4.00
chr20-1684863-1686631	18.00	18.00
chr20-24004981-24007178	8.00	8.00
chr20-61864701-61866469	8.00	8.00
chr20-49308216-49309521	78.00	78.00
chr20-6636612-6639499	27.00	27.00
chr20-51045088-51047468	34.00	34.00
chr20-58243435-58245502	35.00	35.00
chr20-44871574-44872758	34.00	34.00
chr20-32128767-32131166	55.00	55.00
chr20-41212692-41214726	135.00	135.00
chr20-28297698-28299651	22.00	22.00
chr20-51964254-51966821	37.00	37.00
chr20-24417808-24420704	49.00	49.00
chr20-62644015-62645101	3.00	3.00
chr20-17491740-17493367	42.00	42.00
chr20-1688906-1691198	18.00	18.00
chr20-24527127-24528574	29.00	29.00
chr20-52428593-52431117	14.00	14.00
chr20-26321024-26322698	3.00	3.00
chr20-45100354-45103113	66.00	66.00
chr20-44040957-44042285	37.00	37.00
chr20-8172782-8174522	5.00	5.00
chr20-18388054-18390831	17.00	17.00
chr20-46328110-46329817	15.00	15.00
chr20-63202423-63203944	1.00	1.00
chr20-32034568-32036641	31.00	31.00
chr20-16380648-16382791	33.00	33.00
chr20-17324416-17326603	10.00	10.00
chr20-62256391-62259148	85.00	85.00
chr20-26695365-26697570	20.00	20.00
chr20-20397612-20400433	71.00	71.00
chr20-45766019-45767059	16.00	16.00
chr20-10489899-10492143	1.00	1.00
chr20-29482814-29485013	12.00	12.00
chr20-32760142-32762589	37.00	37.00
chr20-48257999-48260762	27.00	27.00
chr20-40817992-40820043	24.00	24.00
chr20-44136060-44138160	58.00	58.00
chr20-52191676-52193518	45.00	45.00
chr20-53054104-53056005	45.00	45.00

Continued on next page

bioRxiv preprint doi: https://doi.org/10.1101/2023.08.02.551686; this version posted October 24, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S5. Comparison of HSPs between BLAT and PxBLAT

Data	BLAT	PxBLAT
chr20-62008822-62010386	1.00	1.00
chr20-56222232-56225186	48.00	48.00
chr20-11993500-11995429	67.00	67.00
chr20-2838968-2840190	2.00	2.00
chr20-44370497-44372521	36.00	36.00
chr20-15651294-15652738	5.00	5.00
chr20-60446354-60447385	1.00	1.00
chr20-53863722-53865382	2.00	2.00
chr20-63911245-63912551	15.00	15.00
chr20-46941439-46942618	21.00	21.00
chr20-38876786-38878888	41.00	41.00
chr20-54691788-54694637	33.00	33.00
chr20-28756677-28757926	2.00	2.00
chr20-46371353-46372831	38.00	38.00
chr20-47942735-47945146	27.00	27.00
chr20-22121418-22122589	4.00	4.00
chr20-18708284-18710520	14.00	14.00
chr20-45981971-45984388	104.00	104.00
chr20-46135966-46137267	23.00	23.00
chr20-1575213-1576967	119.00	119.00
chr20-39187678-39189005	9.00	9.00
chr20-41278984-41281397	151.00	151.00
chr20-9494396-9496948	113.00	113.00
chr20-31727762-31729484	35.00	35.00
chr20-36022112-36023674	57.00	57.00
chr20-48271564-48273408	135.00	135.00
chr20-40669278-40670791	8.00	8.00
chr20-42986488-42988685	10.00	10.00
chr20-1299958-1301679	106.00	106.00
chr20-13502614-13505127	43.00	43.00
chr20-53382502-53385117	197.00	197.00
chr20-52713929-52716436	6.00	6.00
chr20-50921819-50923547	32.00	32.00
chr20-59129918-59131065	9.00	9.00
chr20-24871142-24872155	3.00	3.00
chr20-36621360-36623326	32.00	32.00
chr20-26210260-26212225	61.00	61.00
chr20-26014391-26017116	44.00	44.00
chr20-30326384-30327411	33.00	33.00
chr20-15482840-15484780	4.00	4.00
chr20-22188810-22191531	7.00	7.00
chr20-53189540-53191340	19.00	19.00
chr20-20165110-20167633	6.00	6.00
chr20-50879702-50882451	29.00	29.00
chr20-63776357-63779251	43.00	43.00
chr20-8593094-8594767	15.00	15.00
chr20-13374024-13375459	149.00	149.00
chr20-7629062-7631848	20.00	20.00
chr20-46281701-46284229	78.00	78.00
chr20-20612704-20614333	6.00	6.00
chr20-7789652-7790884	51.00	51.00
chr20-20211163-20212564	14.00	14.00
chr20-16947414-16949607	21.00	21.00
chr20-54513399-54515370	37.00	37.00

Continued on next page

bioRxiv preprint doi: https://doi.org/10.1101/2023.08.02.551686; this version posted October 24, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S5. Comparison of HSPs between BLAT and PxBLAT

Data	BLAT	PxBLAT
chr20-56480693-56482358	31.00	31.00
chr20-45906833-45908007	60.00	60.00
chr20-25103924-25105179	4.00	4.00
chr20-20127481-20130326	65.00	65.00
chr20-64287371-64289651	38.00	38.00
chr20-48414492-48416468	25.00	25.00
chr20-24558966-24560320	27.00	27.00
chr20-51647915-51650334	35.00	35.00
chr20-38255609-38258269	52.00	52.00
chr20-63504410-63506929	16.00	16.00
chr20-54586295-54587370	29.00	29.00
chr20-44482370-44485031	36.00	36.00
chr20-50765756-50767150	54.00	54.00
chr20-366563-367950	15.00	15.00
chr20-63477911-63480561	94.00	94.00
chr20-15045333-15048254	23.00	23.00
chr20-7274439-7276379	172.00	172.00
chr20-59961834-59963681	15.00	15.00
chr20-22119668-22121042	211.00	211.00
chr20-6377058-6378441	193.00	193.00
chr20-54875756-54878074	214.00	214.00
chr20-18239510-18240655	70.00	70.00