



PRINCIPE FELIPE
CENTRO DE INVESTIGACION

Interactive web tool to manage sequencing data for the detection of viral insertion sites in gene therapy experiments



José M. Juanes^{1,2}, Joaquín Tárraga¹, Asunción Gallego¹, Ignacio Medina⁵,
Vicente Arnau^{1,2} and Joaquín Dopazo^{1,3,4}

¹ Computational Genomics Department, Centro de Investigación Príncipe Felipe (CIPF), 46012, Valencia, Spain.

² Departamento de Informática, Escola Tècnica Superior d'Enginyeria, Universitat de València, 46100, Burjassot (Valencia), Spain.

³ Bioinformatics of Rare Diseases (BIER), CIBER de Enfermedades Raras (CIBERER), 46012, Valencia, Spain.

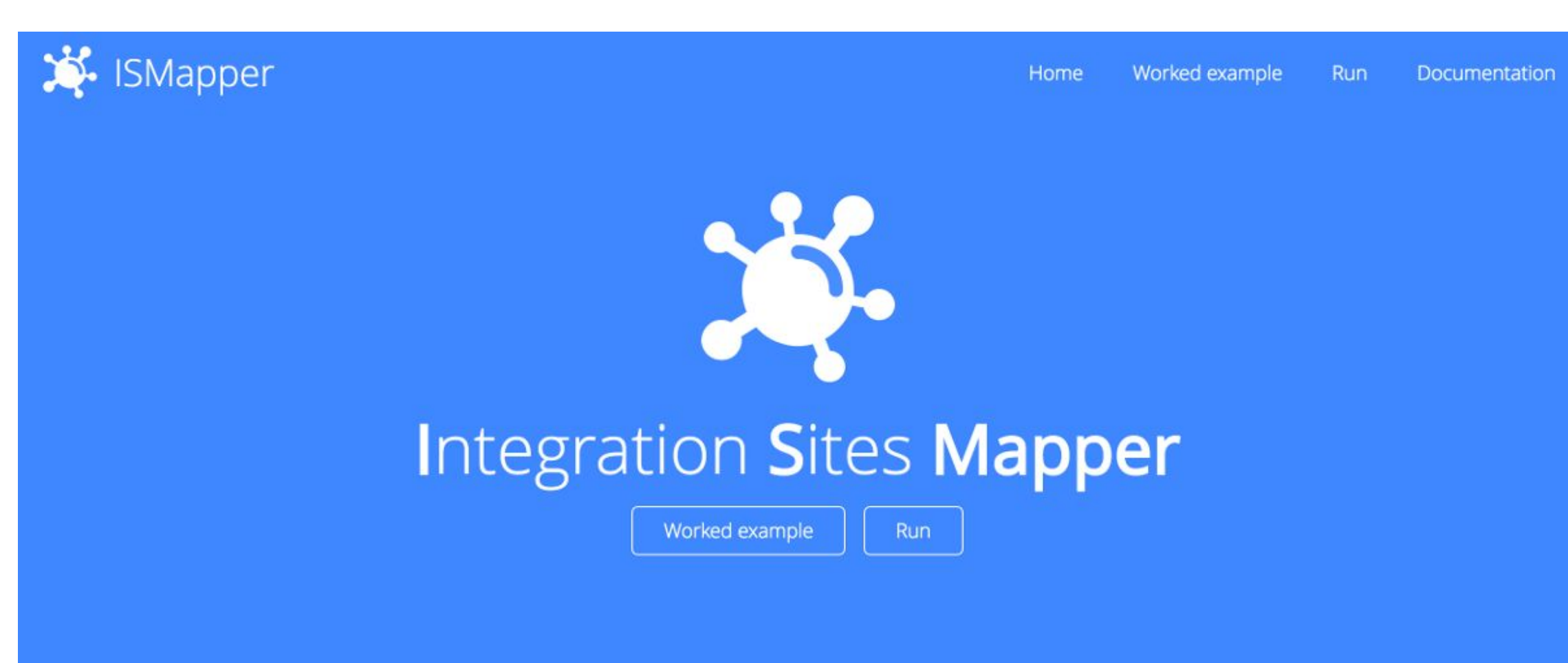
⁴ Functional Genomics Node, (INB, PRB2, ISCIII) at CIPF, Valencia, 46012, Spain.

⁵ HPC Service, University Information Services, University of Cambridge.

INTRODUCTION

The possibility of integrating viral vectors to become a persistent part of the host genome makes them a crucial element of clinical gene therapy. However, viral integration has associated risks, such as the unintentional activation of oncogenes that can derive in cancers. Therefore, the analysis of integration sites of retroviral vectors is a crucial step in developing safer vectors for therapeutic use. Here we present **ISMapper**, a vector integration site analysis web server to analyse next-generation sequencing data for retroviral vector integration sites. Because it uses novel mapping algorithms, **ISMapper** is remarkably faster than previous available options and provides a useful interactive graphical interface to analyse the integration sites found in the genomic context.

WEB APPLICATION <http://ismapper.babelomics.org/>



ISMapper

HomeWorked exampleRunDocumentation

ISMapper

HomeWorked exampleRunDocumentation

Upload and Run

Here you can create a new project by uploading your data files. You can test the tool by uploading our [sample data](#).

Please note that **every project is available up to 60 days**. After this time, the data in our server will be removed, and you must upload again your data if you want to continue using the tool.

Title

You can set a title for your project. Use it for provide a short description of your experiment.

Select files

Press the next button for select your **FASTA** or **FASTQ** file containing the reads of your experiment. Additionally, if your file is too big, you can compress in a **zip** and select it.

Seleccionar archivo

Ningún archivo seleccionado

Select aligner

Select the aligner tool to align the reads:

BWA

Select reference genome

Select the reference genome:

Homo Sapiens GRCh38

Select minimum mapping quality

Select the minimum mapping quality of the reads (default is 20):

20

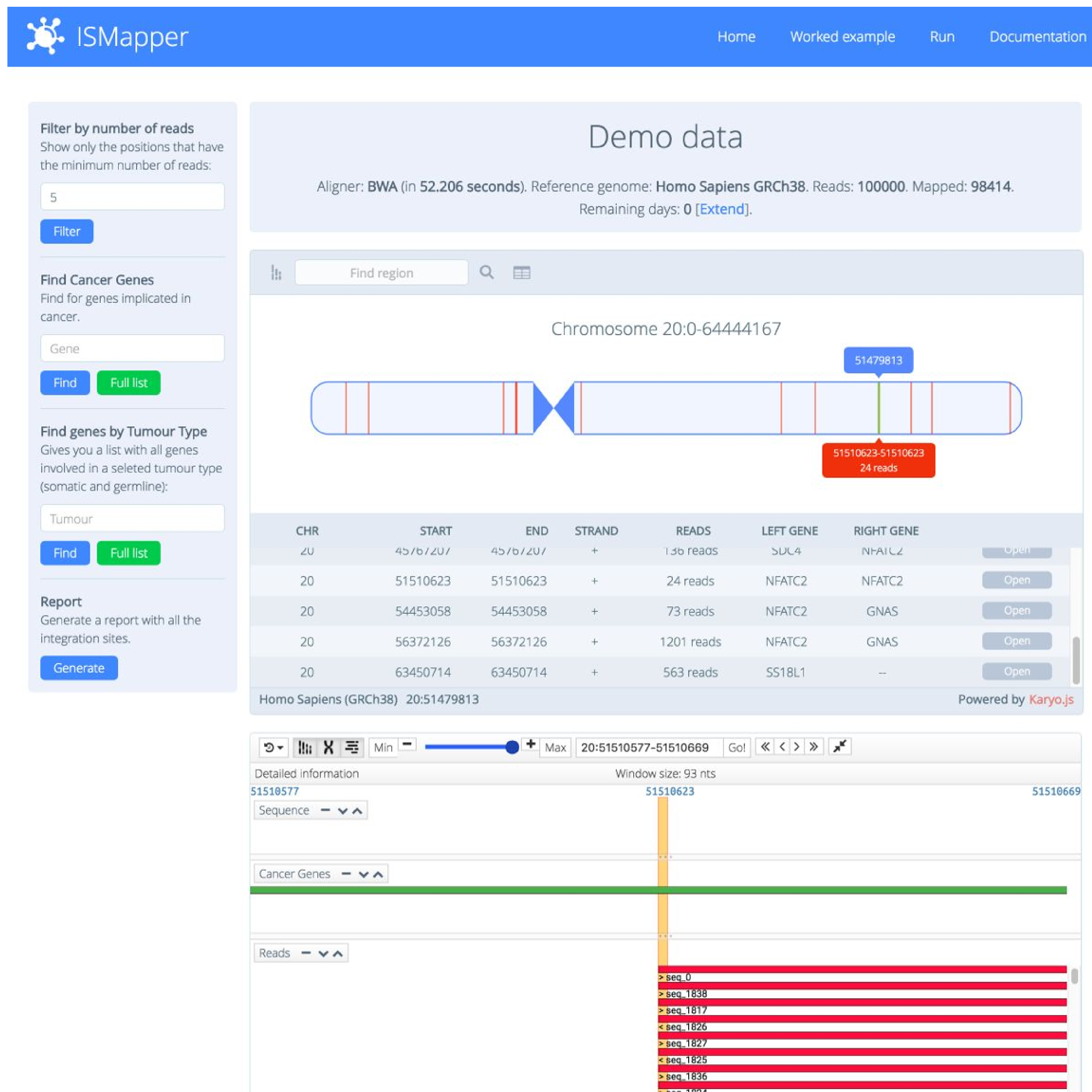
Your email (optional)

You can provide us an email and we will notify you when the process is completed. Also, we will provide you a link where you can visualize your data.

Run your experiment

Check if all the information that you have provided is correct and press the next button for upload your data to the server and start the analysis:

Upload and run



Demo data

Aligner: **BWA** (in **52.206 seconds**). Reference genome: **Homo Sapiens GRCh38**. Reads: **100000**. Mapped: **98414**. Remaining days: **0** [\[Extend\]](#).

Find region

Chromosome 20:0-64444167

51479813

51510623-51510623
24 reads

CHR	START	END	STRAND	READS	LEFT GENE	RIGHT GENE	
20	51510623	51510623	+	24 reads	NFATC2	NFATC2	Open
20	54453058	54453058	+	73 reads	NFATC2	GNAS	Open
20	56372126	56372126	+	1201 reads	NFATC2	GNAS	Open
20	63450714	63450714	+	563 reads	SS18L1	...	Open

Homo Sapiens (GRCh38) 20:51479813

Powered by [Karyo.js](#)

Min

Max

20:51510577-51510669

Go!

<<>>

Detailed information

Window size: 93 nts

51510577

51510623

51510669

Sequence

Cancer Genes

Reads

> seq_0

> seq_1838

> seq_1817

< seq_1826

< seq_1827

< seq_1825

> seq_1836

> seq_1824

REFERENCES

- [1] Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25, 1754-1760.
- [2] Tárraga, J., Arnau, V., Martínez, H., Moreno, R., Cazorla, D., Salavert-Torres, J., Blanquer-Espert, I., Dopazo, J. and Medina, I. (2014) Acceleration of short and long DNA read mapping without loss of accuracy using suffix array. *Bioinformatics*, 30, 3396-3398.
- [3] Medina, I., Salavert, F., Sanchez, R., de Maria, A., Alonso, R., Escobar, P., Bleda, M. and Dopazo, J. (2013) Genome Maps, a new generation genome browser. *Nucleic Acids Res*, 41, W41-46.