SOFTWARE TOOL ARTICLE

REVISED Biomedical Mutation Analysis (BMA): A software tool for analyzing mutations associated with antiviral resistance [version 2; referees: 2 approved]

Karina Salvatierra[1], Hector Florez[2]

[1]Faculty of Exact, Chemical and Natural Sciences, Universidad Nacional de Misiones, Posadas, Argentina
[2]Faculty of Technology, Universidad Distrital Francisco José de Caldas, Bogotá, Colombia

v2 **First published:** 03 Jun 2016, **5**:1141 (doi: 10.12688/f1000research.8740.1)
Latest published: 02 Aug 2016, **5**:1141 (doi: 10.12688/f1000research.8740.2)

Abstract

Introduction: Hepatitis C virus (HCV) is considered a major public health problem, with 200 million people infected worldwide. The treatment for HCV chronic infection with pegylated interferon alpha plus ribavirin inhibitors is unspecific; consequently, the treatment is effective in only 50% of patients infected. This has prompted the development of direct-acting antivirals (DAA) that target virus proteins. These DAA have demonstrated a potent effect *in vitro* and *in vivo*; however, virus mutations associated with the development of resistance have been described.
Objective: To design and develop an online information system for detecting mutations in amino acids known to be implicated in resistance to DAA.
Materials and methods: We have used computer applications, technological tools, standard languages, infrastructure systems and algorithms, to analyze positions associated with resistance to DAA for the NS3, NS5A, and NS5B genes of HCV.
Results: We have designed and developed an online information system named Biomedical Mutation Analysis (BMA), which allows users to calculate changes in nucleotide and amino acid sequences for each selected sequence from conventional Sanger and cloning sequencing using a graphical interface.
Conclusion: BMA quickly, easily and effectively analyzes mutations, including complete documentation and examples. Furthermore, the development of different visualization techniques allows proper interpretation and understanding of the results.
The data obtained using BMA will be useful for the assessment and surveillance of HCV resistance to new antivirals, and for the treatment regimens by selecting those DAA to which the virus is not resistant, avoiding unnecessary treatment failures. The software is available at: http://bma.itiud.org
.

Open Peer Review

Referee Status: ✓ ✓

	Invited Referees 1	2
REVISED version 2 published 02 Aug 2016		✓ report
version 1 published 03 Jun 2016	✓ report	? report

1 **Vicente Perez-Brocal**, Public Health and Epidemiology Network Biomedical Research Center (CIBERESP) Spain

2 **Carla García-Morales**, Universidad Autónoma del Estado de México Mexico

Discuss this article

Comments (0)

Corresponding author: Hector Florez (haflorezf@udistrital.edu.co)

How to cite this article: Salvatierra K and Florez H. **Biomedical Mutation Analysis (BMA): A software tool for analyzing mutations associated with antiviral resistance [version 2; referees: 2 approved]** *F1000Research* 2016, **5**:1141 (doi: 10.12688/f1000research.8740.2)

Grant information: This work has been supported by the Information Technologies Innovation (ITI) Research Group.

Competing interests: No competing interests were disclosed.

First published: 03 Jun 2016, **5**:1141 (doi: 10.12688/f1000research.8740.1)

REVISED **Amendments from Version 1**

1. Information regarding the list of positions has been updated including an explanation of the column "Main".

2. Figure 6 has been also updated in order to present the popup messages of the force directed graph. In addition, the explanation of this figure has been extended.

See referee reports

Introduction

Chronic hepatitis C infection is caused by the hepatitis C virus (HCV) and affects an estimated 200 million people worldwide[1,2]. Transmission occurs by percutaneous exposure through blood products. The major risk factors for HCV infection are parenteral exposure, and needle sharing among intravenous drug users. In addition, hemodialysis patients are at risk of contracting an HCV infection[3–5].

Historically, HCV drug therapy has depended on interferon-α and ribavirin and the effectiveness of this combination therapy are primarily determined by the HCV genotype[6]. The advent of direct-acting antiviral agents (DAA) has paved the way for a new era for the treatment of HCV infection. The most important contribution in their development primarily target protease NS3, protein NS5A or NS5B RNA-dependent RNA polymerase[7]. However, because the emergence of resistant viral variants, DAA is one of the factors to be taken into account in the treatment[8]. Antiviral capacity may be limited by the ability of the virus to develop resistance to new antivirals[9]. Resistance mutations to DAA have been observed both *in vitro* and *in vivo*[10–12]. In addition, people infected with HCV, who are left untreated, can develop natural viral variants harboring resistance mutations. Current data indicates pre-existing mutations to NS3 protease inhibitors, NS5A inhibitors, and non-nucleoside inhibitors of NS5B polymerase in 7.7%, 16.2% and 22.5%, of infected patients[13–16]. Probably these viral variants contribute to the selection of resistance to DAA during the initial weeks of monotherapy[17–20].

Using DAA implies the possibility of selection of resistant variants. Antiviral resistance results from amino acid substitutions that produce conformational changes that interfere with drug-target interaction. These mutations typically involve a biological cost, and viruses carrying these mutations are found in smaller numbers than wild-type viruses; however, they can be positively selected during therapy[21].

Genetic variability affects the response to old and new therapies. It is therefore important to determine mutations of resistance to antiviral drugs.

There is an increasing need to develop bioinformatic tools to analyze the rapidly growing amount of nucleotide and amino acid sequence data in different organisms such as viruses. An important task in bioinformatics is the provisioning of data and tools in a simple manner for users to locate and use. Sequencing generates large amounts of data that need to be analyzed. Advances in information technology have stimulated the development of new computer applications and algorithms for data analysis, and computer visualization tools for the representation of variation patterns. The analysis of mutations is important to understand antiviral resistance and to understand the functions of different proteins. The aim of this study was to develop an online information system named Biomedical Mutation Analysis (BMA), which allows users to calculate changes in nucleotide and amino acid sequences for each selected sequence through a graphical interface.

Materials and methods

To create the online information system, we used different standard tools, languages, and infrastructure systems. BMA was designed using the Unified Modeling Language (UML)[23], which allows describing the system following the Object Oriented Paradigm. Regarding the development of BMA, we used PHP language version 5.3.29 (https://secure.php.net/), which is supported by Apache software version 2.4.7 (http://www.apache.org/) as the application server. For the front end of BMA, we used Bootstrap version 3.3.6 (http://getbootstrap.com/), which is the most popular HTML, CSS, and JavaScript framework for developing responsive web projects. BMA also has some features based on JavaScript language supported by JQuery version 1.12.3 (https://jquery.com/), which is a JavaScript library that facilitates some specific JavaScript functionalities. BMA provides three different outputs, where two of them use additional support. The former result is a report generated as a pdf file, which is built using ezpdf version 0.0.9 (https://github.com/rebuy-de/ezpdf), which is a library that supports the creation of pdf files. The latter result is a force-directed graph, which is created using D3 (Data-Driven Documents) version 3.5.16 (https://d3js.org/), which is an online JavaScript library that helps to deploy data using fancy visualizations.

BMA stores all information related to the mutation analyses in one database supported by MySql version 5.7.12 (https://www.mysql.com/), which is a relational database management system. The database includes the entities and relationships required for handling all information related to the proposed mutation analyses. The database is manipulated through project phpMyAdmin version 4.3.11 (https://www.phpmyadmin.net/), which is software written in PHP intended to handle the administration of data stored in MySql databases.

The database was designed using the tool MySql Workbench version 6.3 (https://www.mysql.com/products/workbench/), while the online system was developed using the tool Eclipse PHP version 3.7.0 (https://eclipse.org/pdt/). BMA is hosted in a Linux Server debian distribution version 8.4, which includes Apache, MySql, and phpMyAdmin for the right operation of BMA.

All software, frameworks, and libraries used in the design and development of BMA have a GNU General Public License (GNU GPL) (http://www.gnu.org/licenses/licenses.en.html), which implies that BMA was completely created using free software.

We used the nucleotide sequence of genes NS3, NS5A and NS5B of Con1 isolated HCV genotype 1b (accession number: AJ238799), extracted from GenBank (www.ncbi.nlm.nih.gov/genbank/) as a reference sequence.

A compilation of resistance mutations previously described *in vivo* and *in vitro* in the literature for the genes NS3, NS5A and NS5B of the HCV were used for computing the number and type of amino acid variants at the corresponding positions associated with resistance to DAA[24,25].

Results

The BMA's core is the analysis algorithm that is able to evaluate multiple patients, where each one can include multiple sequences. In addition, the algorithm can analyze desired positions that the analyst can define. The execution of the algorithm is just one part of the complete analysis process. The analysis process includes the following steps:

1. The analyst accesses BMA via the web site and selects the option "HCV" from the "Mutation Analysis" menu. BMA presents the list of genes available for HCV, which includes the name, description, and reference sequence (by clicking on the corresponding icon). Figure 1 presents the list of available genes.

2. The analyst can use the search icon placed in each gene of the HCV (e.g., NS3, NS5A, NS5B) to proceed to the following step, which corresponds to the selection of the positions to be analyzed. Thus, possible positions are sorted in a list, which includes the number of the position, mutation, antiviral name, inhibitor class, a flag ("Yes" or "No") that indicates whether or not the position is a main position for the selected gene, and references that can be *in vitro* or *in vivo*. It is important to mention that BMA is flexible allowing the inclusion of further positions, mutations, and antivirals established in new or future research. Regarding references, each position presents the list of academic papers that support scientifically the inclusion of the position in the mutation analysis. Furthermore, for each position, there is an icon that lists the reference details with a link that redirects to one academic search service with the information of the selected reference. Figure 2 presents the list of some positions for the gene NS3.

3. After selecting the positions to perform the analysis, the analyst is asked to provide the patient sequences as plain text files. BMA offers an example dataset for testing the analysis. BMA can automatically read and analyze multiple data files sequentially. These data files may contain a varying number of sequences that represent one patient. BMA can recognize plain text files, which must include the symbol '>' and the sequence name in the first line of the file. The sequence data starts on the second line. Nucleotide data must be written in one line. The sequence must include the symbols: A, C, G, T. Sequences can also include the symbol '-' for specifying missing data. In sequences, blank spaces, tabs, break lines and other symbols are not accepted (see Figure 3).

4. Once patient files are selected, the analysis algorithm is executed. The algorithm presents the results in three different ways:

a) Online textual visualization of necessary nucleotide changes that produce an amino acid change, which generates resistance (Figure 4).

b) An automatically generated report, which is sent to the analyst's e-mail address. This report contains a summary of the calculated mutations for each sequence and the full detailed report of the executed analysis (Figure 5).

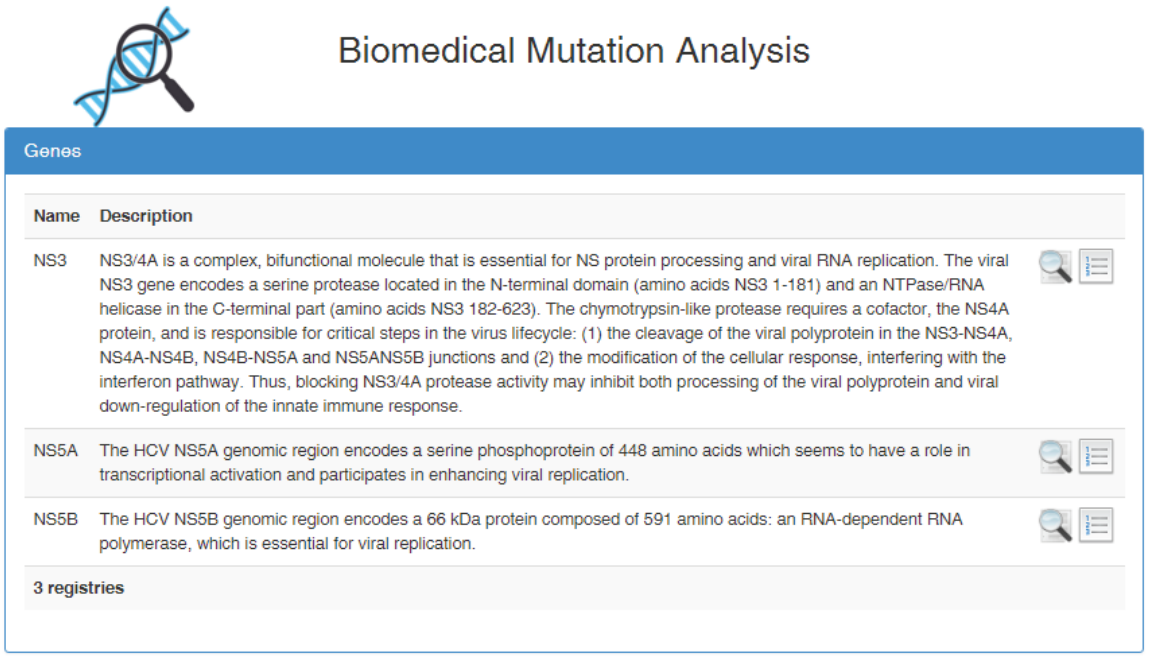
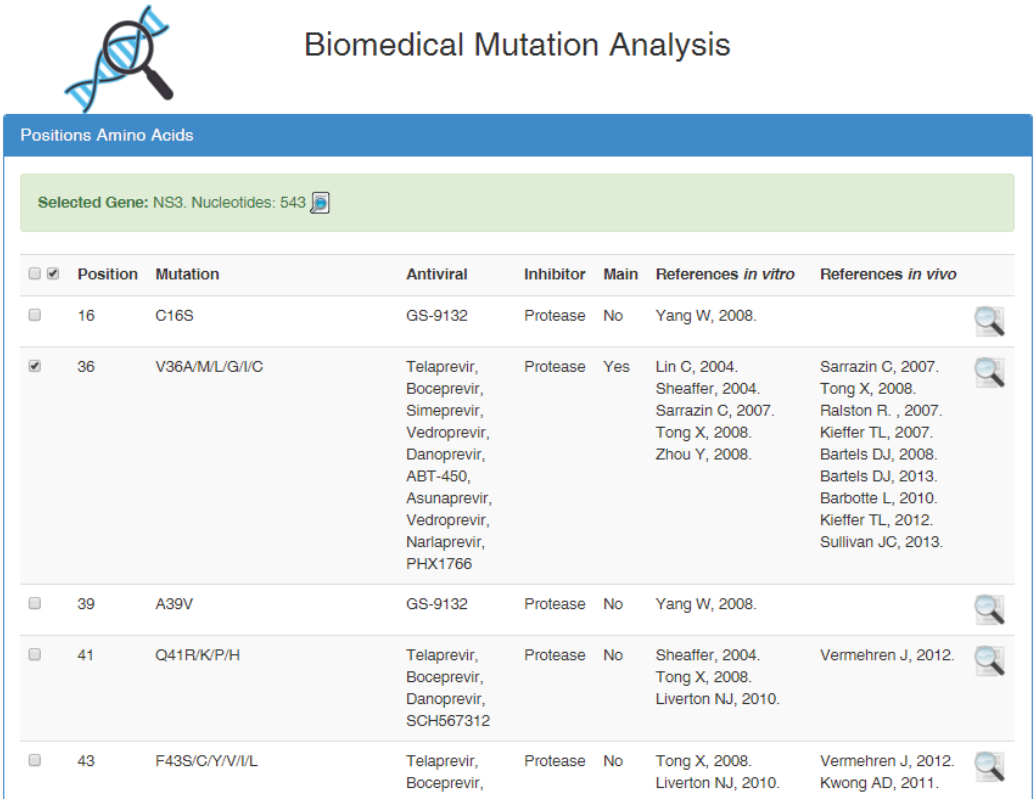


Figure 1. List of available genes for HCV.



Position	Mutation	Antiviral	Inhibitor	Main	References *in vitro*	References *in vivo*
16	C16S	GS-9132	Protease	No	Yang W, 2008.	
36	V36A/M/L/G/I/C	Telaprevir, Boceprevir, Simeprevir, Vedroprevir, Danoprevir, ABT-450, Asunaprevir, Vedroprevir, Narlaprevir, PHX1766	Protease	Yes	Lin C, 2004. Sheaffer, 2004. Sarrazin C, 2007. Tong X, 2008. Zhou Y, 2008.	Sarrazin C, 2007. Tong X, 2008. Ralston R. , 2007. Kieffer TL, 2007. Bartels DJ, 2008. Bartels DJ, 2013. Barbotte L, 2010. Kieffer TL, 2012. Sullivan JC, 2013.
39	A39V	GS-9132	Protease	No	Yang W, 2008.	
41	Q41R/K/P/H	Telaprevir, Boceprevir, Danoprevir, SCH567312	Protease	No	Sheaffer, 2004. Tong X, 2008. Liverton NJ, 2010.	Vermehren J, 2012.
43	F43S/C/Y/V/I/L	Telaprevir, Boceprevir,	Protease	No	Tong X, 2008. Liverton NJ, 2010.	Vermehren J, 2012. Kwong AD, 2011.

Figure 2. List of positions (5 of 26) for the gene NS3.

```
>Patient1
GCGCCTATCACGGCCTACGCCCAACAGACGCGGGGCCTACTTGGCTGCATCATCACCAGCCTCACAGGTCGGGACAAGAACCAGGTCGAGGGAGAGGTTCAAGTGGT
CTCCACTGCAACACAATCTTTCCTGGCGACCTGTGTCAACGGCGTGTGTTGGACTGTTTTCCACGGCGCCGGCTCTAAGACCCTGGCCGGCTCAAAAGGCCCAATCA
CTCAAATGTACACCAATGTAGATCAAGACCTCGTCGGTTGGCAGGCCCCTCCAGGGGCGCGTTCTTTGACACCATGCACCTGTGGTAGCTCAGACCTTTACTTGGTC
ACGAGGCATGCTGATGTCATCCCGGTACGCCGGCGAGGCGACAGCAGGGGGAGCCTGCTCTCCCCCAGGCCTGTCTCCTACTTGAAGGGCTCTTCGGGCGGTCCACT
GCTCTGCCCTTCGGGGCACGCCGTAGGCATCTTCCGAGCTGCTGTGTGCACCCGGGGGGTTGCGAAGGCGGTGGACTTCATACCCGTTGAGTCTATGGAAACTACCA
TGCGGTCC
>sequence1
GCGCCTATCACGGCCTACGCCCAACAGACGCGGGGCCTACTTGGCTGCATCATCACCAGCCTCACAGGTCGGGACAAGAACCAGGTCGAGGGAGAGGTTCAAGTGGT
CTCCACTGCAACACAATCTTTCCTGGCGACCTGTGTCAACGGCGTGTGTTGGACTGTTTTCCACGGCGCCGGCTCTAAGACCCTGGCCGGCCCAAAAGGCCCAATCA
CTCAAATGTACACCAATGTAGATCAAGACCTCGTCGGTTGGCAGGCCCCTCCAGGGGCGCGGTCTTTGACACCATGCACCTGTGGTAGCTCAGACCTTTACTTGGTC
ACGAGGCATGCTGATGTCATCCCGGTACGCCGGCGAGGCGACAGCAGGGGGAGCCTGCTCTCCCCCAGGCCTGTCTCCTACTTGAAGGGCTCTTCGGGCGGTCCACT
GCTCTGCCCATCGGGGCACGCCGTAGGCATCTTCCGAGCTGCTGTGTGCACCCGGGGGGTTGCGAAGGCGGTGGACTTCATACCCGTTGAGTCTATGGAAACTACCA
TGCGGTCC
>sequence2
GCGCCTATCACGGCCTACGCCCAACAGACGCGGGGCCTATTTGGCTGCATCATCACCAGCCTCACAGGTCGGGACAAGAACCAGGTCGAGGGAGAGGTTCAAGTGGT
CTCCACTGCAACACAATCTTTCCTGGCGACCTGTGTCAACGGCGTGTGTTGGACTGTTTTCCACGGCGCCGGCTCTAAGACCCTGGCCGGCTCAAAAGGCCCAATCA
CTCAAATGTACACCAACGTAGATCAAGACCTCGTCGGTTGGCAGGCCCCTCCAGGGGCGCGTTCTTTGACACCATGCACCTGTGGTAGCTCAGACCTTTACTTGGTC
ACGAGGCATGCTGATGTCATCCCGGTACGCCGGCGAGGCGACAGCAGGGGGAGCCTGCTCTCCCCCAGGCCTGTCTCCTACTTGAAGGGCTCTTCGGGCGGTCCACT
GCTCTGCCCTTCGGGGCACGCCGTAGGCATCTTCCGAGCTGCTGTGTGCACCCGGGGGGTTGCGAAGGCGGTGGACTTCATACCCGTTGAGTCTATGGAAACTACCA
TGCGGTCC
```

Figure 3. Patient file format.

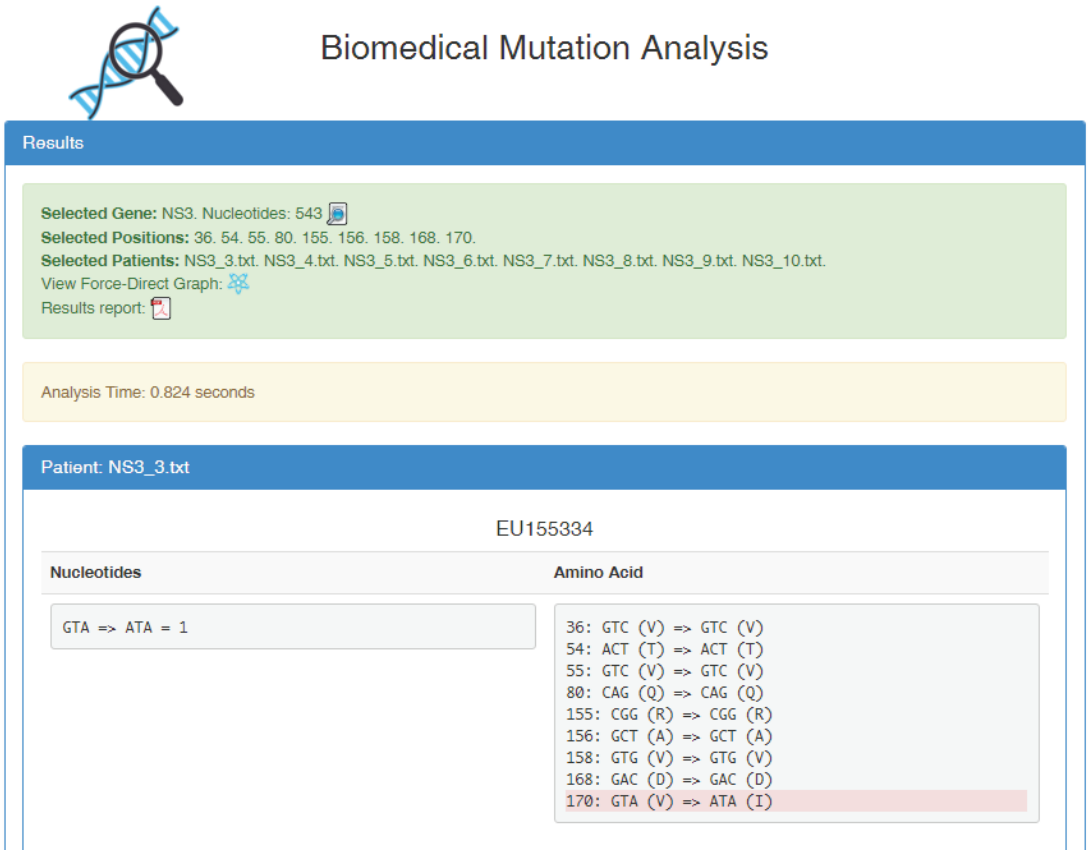


Figure 4. Online textual visualization.

c) A "force-directed" graph that identifies mutations of each patient sequence through node grouping, which corresponds to each analyzed sequence. Figure 6 shows that nodes with the same color corresponds to the sequences of a patient. Nodes of a patient are grouped depending on the amount of mutations. In order to know which sequences have mutated, the analyst can place the mouse pointer over a node and the graph presents a popup message with the information of the corresponding node. For instance, Figure 6 also presents a popup message with the information of the node above it. It indicates three facts: 1) the node corresponds to the patient of the file "NS3_7.txt", 2) the node corresponds to the sequence "FJ864759", and 3) the sequence has at the position 170 one mutation because in this position the aminoacid should be "GTA(V)", but it is "ATA(I)".

For reliable calculations the sequences must contain a substantial part of the genes NS3, NS5A or NS5B (amino acids, aa 1–200).

The analysis algorithm is based on multiple iterations. It collects all patients' plain text files and iterates in order to analyze all of them independently. For each plain text file, the algorithm collects all sequences. Later on, for each sequence, the algorithm performs a new iteration using the selected positions. Then, for each position, it compares the nucleotide and amino acid of the iterated patient sequence with the reference sequence in the iterated position. At this stage, the information about changes is collected with the corresponding patient, sequence, and position. By finishing the execution of the algorithm, BMA uses the collected results to provide the three aforementioned visualizations.

It is important to mention that BMA cannot align sequences. There are some programs that can do this. For example, CLUSTAL W[26] allows multiple alignments. In addition, DNA sequences cannot be edited or manipulated by BMA. No clinical decision should be based only on the result of BMA.

BIOMEDICAL MUTATION ANALYSIS
ANALYSIS REPORT
Date of Analysis: 2015-11-11
Gene: NS3 Con-1(1b)

SUMMARY
Evaluated Positions: 36. 54. 55. 80. 155. 156. 158. 168. 170.

Patient File Name	Sequences	Nucleotide Changes	Amino Acid Changes
NS3_3.txt	EU155334	1	1
	EU155356	1	0
	EF407501	2	0
	KC123699	1	1
	FJ864741	2	0
	EU256000	2	0
	KC123973	5	1
	KC124611	5	1
	KC124112	2	0
	JQ253541	2	0
	FJ864737	3	0
	KC123893	3	1
	JQ246524	3	1
	FJ864774	1	0
	EU155360	3	1
	EU155217	2	1
	KC123999	6	0
	KC124025	2	1
	AB049095	3	1
	KC124485	3	1

Figure 5. **Analysis report.**

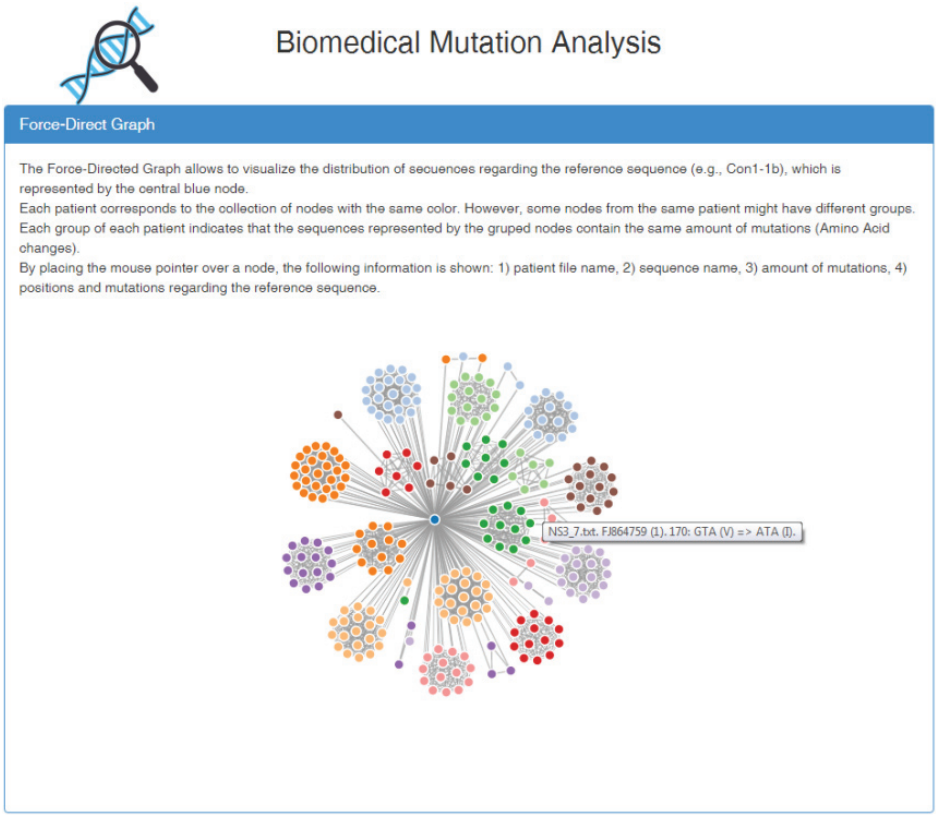


Figure 6. **Force-Directed Graph visualization.**

Conclusions

Software for the detection of mutations associated with resistance to new DAAs is an important tool, because it guarantees accurate and reliable results. Moreover, BMA is freely available, which is different from others such as Bioedit, VectorNTI or MEGA, because it not only allows researchers to perform analysis for the identification of mutations, but also provides detailed information of mutations' positions, amino acid changes as well as antiviral information and related literature of resistance mutations to the DAA. When BMA is compared with other available tools (e.g., HCV.geno2pheno), it is different because it provides details of the nucleotides changes that produce an amino acid change.

We obtained an online information system "BMA" that was designed and developed, for performing mutation analysis. BMA provides a suitable analysis facilitating all data management. The results can be visualized in a text report as well as graphically.

BMA provides a quick, easy, and effective computer-based analysis of mutations, including complete documentation and examples. Furthermore, the development of different visualization techniques allows for proper interpretation and understanding of the results. The data obtained by BMA will be useful for the assessment and surveillance of HCV resistance to new antivirals, and for the treatment regimens by selecting those DAAs to which the virus is not resistant, avoiding unnecessary treatment failures.

BMA has been designed to be flexible and adaptable. It is a great advantage because it can be used for future evaluation of other viruses such as Influenza and even microorganisms such as bacteria or parasites. Thus, as future work, BMA will analyze a wide range of pathogens. In addition, BMA might be upgraded in order to offer new visualization techniques for facilitating the interpretation of the obtained analysis.

BMA has a small disadvantage. It requires a specific format of sequence information, which is very similar to the FASTA format; thus, the preparation of such information might require a small additional effort. In addition, in future versions, BMA will accept different file formats such as the FASTA format.

Software availability

Software available from: http://bma.itiud.org

Latest source code: https://github.com/florezfernandez/bma

Archived source code as at the time of publication: http://dx.doi.org/10.5281/zenodo.50994[27]

License: GNU General Public License (GPL)

Author contributions

Karina Salvatierra reviewed the literature and wrote the manuscript. Hector Florez designed and developed the BMA software and wrote the manuscript.

Competing interests

No competing interests were disclosed.

Grant information

This work has been supported by the Information Technologies Innovation (ITI) Research Group.

Acknowledgments

The authors would like to thank the ITI Research Group that belongs to the Universidad Distrital Francisco José de Caldas (Colombia), and Professor Jorge E. Osorio, Department of Pathobiological Sciences, University of Wisconsin-Madison (USA), for its collaboration in the project, making this research possible.

References

1. Hajarizadeh B, Grebely J, Dore GJ: **Epidemiology and natural history of HCV infection.** *Nat Rev Gastroenterol Hepatol.* 2013; **10**(9): 553–62.
PubMed Abstract | Publisher Full Text

2. Mohd Hanafiah K, Groeger J, Flaxman AD, *et al.*: **Global epidemiology of hepatitis C virus infection: new estimates of age-specific antibody to HCV seroprevalence.** *Hepatology.* 2013; **57**(4): 1333–42.
PubMed Abstract | Publisher Full Text

3. Alter MJ: **HCV routes of transmission: what goes around comes around.** *Semin Liver Dis.* 2011; **31**(4): 340–6.
PubMed Abstract | Publisher Full Text

4. Pondé RA: **Hidden hazards of HCV transmission.** *Med Microbiol Immunol.* 2011; **200**(1): 7–11.
PubMed Abstract | Publisher Full Text

5. Salvatierra K, Florez H: **Analysis of hepatitis C virus in hemodialysis patients.** *Infectio.* 2015.
Publisher Full Text

6. McHutchison JG, Lawitz EJ, Shiffman ML, *et al.*: **Peginterferon alfa-2b or alfa-2a with ribavirin for treatment of hepatitis C infection.** *N Engl J Med.* 2009; **361**(6): 580–93.
PubMed Abstract | Publisher Full Text

7. Pawlotsky JM: **Treatment of chronic hepatitis C: current and future.** *Curr Top Microbiol Immunol.* 2013; **369**: 321–42.
PubMed Abstract | Publisher Full Text

8. Cortez KJ, Maldarelli F: **Clinical management of HIV drug resistance.** *Viruses.* 2011; **3**(4): 347–78.
PubMed Abstract | Publisher Full Text | Free Full Text

9. Salvatierra K, Fareleski S, Forcada A, *et al.*: **Hepatitis C virus resistance to new specifically-targeted antiviral therapy: A public health perspective.** *World J Virol.* 2013; **2**(1): 6–15.
PubMed Abstract | Publisher Full Text | Free Full Text

10. Kwong AD, McNair L, Jacobson I, *et al.*: **Recent progress in the development of selected hepatitis C virus NS3.4A protease and NS5B polymerase inhibitors.** *Curr Opin Pharmacol.* 2008; **8**(5): 522–31.
PubMed Abstract | Publisher Full Text

11. Welsch C, Domingues FS, Susser S, *et al.*: **Molecular basis of telaprevir resistance due to V36 and T54 mutations in the NS3-4A protease of the hepatitis C virus.** *Genome Biol.* 2008; **9**(1): R16.
PubMed Abstract | Publisher Full Text | Free Full Text

12. Wyles DL: **Antiviral resistance and the future landscape of hepatitis C virus infection therapy.** *J Infect Dis.* 2013; **207**(Suppl 1): S33–9.
PubMed Abstract | Publisher Full Text

13. López-Labrador FX, Moya A, Gonzàlez-Candelas F: **Mapping natural polymorphisms of hepatitis C virus NS3/4A protease and antiviral resistance to inhibitors in worldwide isolates.** *Antivir Ther.* 2008; **13**(4): 481–94.
PubMed Abstract

14. Margeridon S, Le Pogam S, Liu T, *et al.*: **Ultra-deep sequencing of the NS3 and**

NS5B regions detects pre-existing resistant variants to direct acting antivirals (DAA) in HCV genotype 1 treatment-naive infected patients. *Hepatology.* 2010; **52**(Suppl 1): 714A.
Reference Source

15. Margeridon S, Le Pogam S, Liu TF, *et al.*: **No detection of variants bearing NS5B S282T mericitabine (MCB) resistance mutation in DAA treatment-naive HCV genotype 1 infected patients using ultra-deep pyrosequencing (UDPS).** *Hepatology.* 2011; **54**(Suppl 1): 532A.
Reference Source

16. Wyles DL, Gutierrez JA: **Importance of HCV genotype 1 subtypes for drug resistance and response to therapy.** *J Viral Hepat.* 2014; **21**(4): 229–40.
PubMed Abstract | Publisher Full Text

17. Sarrazin C, Kieffer TL, Bartels D, *et al.*: **Dynamic hepatitis C virus genotypic and phenotypic changes in patients treated with the protease inhibitor telaprevir.** *Gastroenterology.* 2007; **132**(5): 1767–77.
PubMed Abstract | Publisher Full Text

18. Sarrazin C, Rouzier R, Wagner F, *et al.*: **SCH 503034, a novel hepatitis C virus protease inhibitor, plus pegylated interferon alpha-2b for genotype 1 nonresponders.** *Gastroenterology.* 2007; **132**(4): 1270–8.
PubMed Abstract | Publisher Full Text

19. Kuntzen T, Timm J, Berical A, *et al.*: **Naturally occurring dominant resistance mutations to hepatitis C virus protease and polymerase inhibitors in treatment-naïve patients.** *Hepatology.* 2008; **48**(6): 1769–78.
PubMed Abstract | Publisher Full Text | Free Full Text

20. Susser S, Vermehren J, Forestier N, *et al.*: **Analysis of long-term persistence of resistance mutations within the hepatitis C virus NS3 protease after treatment with telaprevir or boceprevir.** *J Clin Virol.* 2011; **52**(4): 321–7.
PubMed Abstract | Publisher Full Text

21. Thompson AJ, McHutchison JG: **Antiviral resistance and specifically targeted therapy for HCV (STAT-C).** *J Viral Hepat.* 2009; **16**(6): 377–87.
PubMed Abstract | Publisher Full Text

22. Excoffier L, Heckel G: **Computer programs for population genetics data analysis: a survival guide.** *Nat Rev Genet.* 2006; **7**(10): 745–58.
PubMed Abstract | Publisher Full Text

23. Rumbaugh J, Jacobson I, Booch G: **The Unified Modeling Language Reference Manual.** Pearson Higher Education. 2004.
Reference Source

24. Salvatierra KA: **Resistencias a nuevos antivirales de acción directa en aislados clínicos del virus de la hepatitis C.** [Ph.D. thesis]. Valencia. Universitat de Valencia, Spain. 2014.
Reference Source

25. Patiño-Galindo JA, Salvatierra K, González-Candelas F, *et al.*: **Comprehensive Screening for Naturally Occurring Hepatitis C Virus Resistance to Direct-Acting Antivirals in the NS3, NS5A, and NS5B Genes in Worldwide Isolates of Viral Genotypes 1 to 6.** *Antimicrob Agents Chemother.* 2016; **60**(4): 2402–16.
PubMed Abstract | Publisher Full Text | Free Full Text

26. Thompson JD, Higgins DG, Gibson TJ: **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.** *Nucleic Acids Res.* 1994; **22**(22): 4673–80.
PubMed Abstract | Publisher Full Text | Free Full Text

27. Florez H, Salvatierra K: **BMA.** *Zenodo.* 2016.
Publisher Full Text

Open Peer Review

Current Referee Status: ✓ ✓

Version 2

Referee Report 12 August 2016

doi:10.5256/f1000research.10090.r15382

Carla García-Morales
Universidad Autónoma del Estado de México, Toluca, Mexico

I have review the new version that the authors of the BMA software have resubmitted. The authors have addressed all my previous comments and I think the tool and the manuscript are appropriate for indexing. All the best.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Competing Interests: No competing interests were disclosed.

Version 1

Referee Report 28 July 2016

doi:10.5256/f1000research.9404.r14899

Carla García-Morales
Universidad Autónoma del Estado de México, Toluca, Mexico

Authors of the BMA tool propose the creation of a bioinformatics tool that identifies changes in nucleotide and amino acid sequences which could represent an impact on antiviral resistance. The suggested system is a promising idea that could be useful for biomedical and pharmacological areas as there is no freely available similar tools.

One of my concerns, if not the main is English vocabulary and grammar, I would suggest the authors to take advantage of an editing service as some statements could be improved. For a start, from the BMA home page, “In biology, a mutation is a change of the nucleotide sequence of the genome of an org…Mutations play a part in both normal and abnormal biological processes including: evolution, cancer, and the development of the immune system, including junction diversity…Mutation can result in several different types of change in sequences. Mutations in genes can have no effect, alter the product of a gene, or prevent the gene from functioning properly or completely.” I would suggest “In biology, a mutation is a change in the nucleotide sequence on the genome of an org…Mutations play an important role in both normal and abnormal biological processes such as: evolution, cancer, and the development

of the immune system, including junction diversity. Mutations can result in several different types of sequence changes; such changes could have no effect, alter the product of the gene, or prevent it from functioning properly or completely."

It would also be advisable that authors present a tutorial link in the BMA site.

As for the manuscript, authors should explain for figure 2 what is shown under the column named "Main"; and again, be careful about expressions such as scientifically support.

As for any publication on bioinformatics tools, F1000 ask the author to provide sufficient details of codes used for the implementation and operation, unfortunately could not find this information; The only code needed to insert different sequences is the > symbol as in other bioinformatics programs, however the authors should be very clear about this, for example, in the 4th line of number 3 in Results, they only mention "BMA can recognize plain text files, but they have to follow a specific format", so what do they mean with specific format?, this line is not needed as they explain the required format immediately. Another point that should be improved is in Figure 3. Introduce the patient sequence in a text file format instead of "plain text"

I would also suggest to change figure 6, for one where the pointer is over the node so that the reader will be able to see what each node represents.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Competing Interests: No competing interests were disclosed.

Author Response 29 Jul 2016

karina salvatierra, Universidad Nacional de Misiones, Argentina

Dear Carla Garcia-Morales.
Thank you so much for your feedback. We want to inform you:

1. We have updated the text in bma.itiud.org as you suggested.
2. In bma.itiud.org, there is in the menu a link called Screencast. It is a video tutorial that presents how to use BMA
3. Information regarding the list of positions has been updating including an explanation of the column "Main"
4. In the section "Software availability", we provide all details about the source code of the project. Indeed, the project is open source, so everyone can download it.
5. Figure 6 has been also updated in order to present the popup messages of the force directed graph. In addition, the explanation of this figure has been extended.

Thank you so much
Best Regards

Competing Interests: Competing Interests: No competing interests were disclosed.

Referee Report 15 July 2016

doi:10.5256/f1000research.9404.r14898

Vicente Perez-Brocal
Public Health and Epidemiology Network Biomedical Research Center (CIBERESP), Madrid, Spain

The article "Biomedical mutation Analysis (BMA): A software tool for analyzing mutations associated with antiviral resistance" describes a new online tool intended for detection of mutations in amino acids implicated in resistance to direct-acting antivirals.

Despite the complexity and high degree of specifications reflected in the manuscript, especially in the material and methods section, the results section describes each step in a straightforward way that facilitates the implementation of this software. The usage of the language makes this article accessible even for non-experts in the field.

The methodology has been validated using the nucleotide sequence of some gene from HCV genotype 1b, and a compilation of resistance mutations previously described in the literature for those genes.

Minor changes:

- In Materials and Methods section, third paragraph, begins as " The database was design using the tool" and should say "The database was design**ed** using the tool".
- Sentence "For reliable calculations the sequences must contain a substantial part of the genes NS3, NS5A or NS5B" in the Results section results quite vague. How much is a substantial part?
- AAD, that appears in the first paragraph of the conclusions, should be replaced with DAA, for congruence with the rest of the text.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Competing Interests: No competing interests were disclosed.

Author Response 29 Jul 2016

karina salvatierra, Universidad Nacional de Misiones, Argentina

Dear Vicente Perez-Brocal

Thank you so much for your feedback.

Best Regards

Competing Interests: Competing Interests: No competing interests were disclosed.